

Acoustic Localization of Puck Impacts on an Ice Hockey Goal Wall using Sound Source Localization Algorithms

Bachelor Thesis

In partial fulfillment of the requirements for the degree

"Bachelor of Science in Engineering"

Study program:

Medical, Health and Sports Engineering

Management Center Innsbruck

Supervisor:

FH-Prof. Bernhard Hollaus, PhD

Author:

Sebastian Seelos

52213502

Declaration in Lieu of Oath

„I hereby declare, under oath, that this bachelor thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extend of the original quotes is indicated.“

Innsbruck, 20.7.25

Place, Date

A handwritten signature consisting of stylized letters, likely 'S.', followed by a horizontal line.

Signature

In this thesis ChatGPT 4.1 was used to paraphrase and smoothen sentence structure. Additional AI, such as Github Copilot supported debugging processes while programming.

Acknowledgement

I wanna say thank you to my girlfriend, who always supported me during the time of writing my thesis. Despite having a lot going on with internship, thesis, job applications and long distance relationship we managed it together through that time. With submitting the thesis and graduating the Bachelor's program at the MCI a new chapter of my life in Amsterdam can finally begin. Thank you to my friends group Wunderbärchen, which made the time in Innsbruck so unique. The study program here only takes three years, but i have made friends for a lifetime. I really value all the time we could spend together, either studying in uni or skiing in the mountains around it. A special Thank you is dedicated to my supervisor FH-Prof. Bernhard Hollaus, PhD. With his expertise and supervising I always had critical support during the last months to fulfill this project.

Kurzfassung

Diese Arbeit präsentiert das Design, Entwicklung und Bewertung eines akustischen Verfahrens zur Schallquellenlokalisierung, um den Auftreffpunkt eines Eishockey Pucks nach einem Schuss auf eine vertikale Wand zu bestimmen. Dies soll die Grundlage für die Weiterentwicklung eines intelligenten Trainingsgeräts bilden, welches die Schussgenauigkeit und das Timing im Eishockey verbessern kann. Das System basiert auf drei analogen Mikrofonen, die hinter einer Holzplatte montiert sind, und nutzt Kreuzkorrelation, Ankunftszeitdifferenz (TDOA) sowie Winkelabschätzung (AOA), um über die Schallquelle der Einschläge in der zweidimensionalen Ebene der Holzplatte entlang der Torlinie zu lokalisieren. Ein Teensy 4.1 Mikrocontroller dient zur synchronen Erfassung der analogen Signale, während die Signalverarbeitung und Lokalisierungsalgorithmen in Python umgesetzt wurden. Die Entwicklung erfolgte in vier aufeinanderfolgenden Stufen, in denen Hardware-Dämpfung, Mikrofonpositionierung und Signalverarbeitung kontinuierlich verbessert wurden. Die Ergebnisse zweier Testreihen zeigen eine stetige Reduktion des Lokalisierungsfehlers. Der finale Prototyp erreicht eine mittlere Genauigkeit von 23 cm. Das System bleibt durch Faktoren wie die automatische Verstärkungsregelung (AGC) der Mikrofone, die Varianz der Signale durch unterschiedliche Schussgeschwindigkeiten sowie Artefakte durch Echos und Reflexionen begrenzt. Der zweite Teil der Arbeit diskutiert die Limitationen des Systems kritisch und vergleicht diese mit aktuellen Ansätzen aus der Literatur. Potenzielle Verbesserungsmöglichkeiten umfassen die Integration von MEMS-Mikrofonen, adaptive Abtastraten sowie Deep-Learning-Methoden zur verbesserten Peak-Detektion und präziseren Kreuzkorrelation. Diese Arbeit bildet somit die Grundlage für ein kostengünstiges, kamerafreies und intelligentes Trainingsgerät, das sich perspektivisch für Echtzeit-Feedback, interaktive Mini-spiele und Anwendungen im Eishockeytraining einsetzen lässt.

Schlagwörter: Schallquellenlokalisierung, Ankunftszeitdifferenz, akustische Winkelbestimmung, Eishockey, Schussgenauigkeit

Abstract

This thesis presents the design, development, and evaluation of an acoustic sound source localization approach. The objective is to accurately determine impact location of an ice hockey puck after being shot at a vertical wall. The developed method serves as the foundation for the further development of an intelligent training device aiming to improve shot accuracy and timing in ice hockey. The system relies on three analog microphones mounted behind a wooden wall, using cross correlation, Time Difference of Arrival and Angle of Arrival principles to localize sound events in the two dimensional plane along the goal line. A Teensy 4.1 microcontroller is used for synchronized analog signal capturing, while signal processing and localization algorithms are implemented with Python. The development exceeded across four stages, with improvements in hardware damping, microphone placement, and signal processing. Experimental results of two testing series show a steady reduction in localization error, with the final prototype achieving a median accuracy of 23 cm. The system remains limited by factors such as automatic gain control (AGC) on the microphones, signal variability due to the differing velocities of the shots, and artifacts of echoes and reflections. The second part of the thesis discusses the system's limitations critically and compares them with state of the art solutions. Opportunities for future improvements include the integration of MEMS microphones, adaptive sampling strategies, and deep learning methods for enhanced peak detection and more accurate cross correlation. This study lays the groundwork for a cost effective, camera-free and smart trainings device, potentially adaptable for real-time training feedback, interactive mini games in the field of ice hockey.

Keywords: Sound Source Localization, Time Difference of Arrival, Angle of Arrival, Ice Hockey, Shot Accuracy

Contents

| | |
|---|-------------|
| 1. Introduction | 1 |
| 1.1. Motivation and Problem Statement | 1 |
| 1.2. Objective and Purpose | 2 |
| 1.3. Sound Source Localization Techniques and Algorithms | 3 |
| 2. Materials and Methodologies | 4 |
| 2.1. Hardware Setup | 4 |
| 2.1.1. Construction and Sensor Integration at the Goal Wall | 4 |
| 2.1.2. Microphone Framework | 7 |
| 2.1.3. Microcontroller and Microphone Communication | 9 |
| 2.2. Software Setup | 11 |
| 2.2.1. Code Architecture | 11 |
| 2.2.2. Capturing Audio Data | 11 |
| 2.2.3. Signal Filtering and Processing | 12 |
| 2.2.4. Calculation Methodology and Location Determination | 16 |
| 2.3. System Testing and Data Recording | 19 |
| 3. Results and Analysis | 24 |
| 3.1. Localization Accuracy Development | 25 |
| 3.2. Spatial Error Distribution | 26 |
| 4. Discussion | 28 |
| 4.1. Evaluation of System Development Stages | 28 |
| 4.2. Future Improvements and Comparison to State of the Art | 30 |
| 5. Conclusion | 33 |
| Bibliography | VII |
| List of Figures | VIII |
| List of Tables | IX |
| List of Symbols | X |
| Abbreviations | XI |
| A. Erster Anhang | XII |

1. Introduction

1.1. Motivation and Problem Statement

Video Assistant referee, Hawk-Eye technology, or Smart Wearables in Athletic Training. Digital devices have gained an increasingly impact and even reformed sports in recent years. Critical and game essential decisions get digital support, eliminating crucial wrong decisions. The famous "Wembley Goal" during the 1966 FIFA World Cup Final, which secured England's first-ever world title, would likely have been judged differently with today's goal-line technology. Technological advancements not only contribute to clear decision making and fair outcomes, but also elevate training or rehabilitation processes to another level. This results in more efficient and accelerated development of athlete performance in many types of sports.

Given the rapid digital transformation, also in sports, it is crucial to keep up and combine the values of traditional training with modern technologies. Especially in terms of efficient progress, AI-supported or sensor-based training and the possibility of real-time analysis are highly beneficial for athletes, sport teams, or organizations [1]. Ice hockey serves as an excellent example of training reformation with smart support, as its fast pace and need for precision benefit from accurate, real-time feedback that smart systems can provide. Wearable devices such as heart rate belts and Intertial Measurement Units (IMU) provide real-time insights into player movement, including biometric data, physical load and fatigue [2]. With video-based performance analysis and emerging virtual training tools, athletes can optimize both their physical and cognitive performance. It allows athletes to simulate game situations and improve their decision-making on the ice [3]. Machine learning applied to IMU data embedded in hockey gloves has shown high accuracy in classifying optimal stick properties, such as flex, blade pattern, and kick point, based on hand kinematics during shooting, demonstrating the potential of Wearables for individualized equipment fitting [4]. In general, coaches and players benefit from performance analytics and smart tools to specialize training sessions and track individual progress. Despite this rapidly evolving integration of modern technologies in ice hockey, there is still a lack of smart training possibilities when it comes to dedicated tools that facilitate the enhancement and analysis of shot accuracy.

Shot accuracy plays a crucial role in ice hockey, where players must react quickly and instinctively to constantly changing situations. Unlike in sports such as soccer or basketball, there is rarely time to prepare and aim precisely before shooting. The puck is small, the game is fast, and scenarios are often unpredictable. Training opportunities for such high-pressure moments are limited, but a smart goal wall, with features

like mini-games or moving targets that appear just before a shot, could offer innovative ways to improve this specific skill.

1.2. Objective and Purpose

The present thesis sought to address this gap by developing a smart training device designed to improve shot accuracy in ice hockey. In addition to supporting individual skill development, the device aims to provide meaningful data-driven feedback. The primary requirement for such a device is a reliable method of tracking the puck's impact locations after a shot on goal. Sports performance tracking technology is evolving rapidly, and moving from basic wearables to real-time, highly accurate data collection containing advanced wireless solutions [5]. However, locating methods in the specific context of a puck shot-on-goal remain limited, due to the puck's small size and its high shot velocity. For instance, Hawk-Eye is a highly accurate and reliable ball-tracking system using advanced image analysis and multiple high-speed cameras. Nevertheless, its complex setup and high cost make it unsuitable for smaller applications or budget-constrained projects [6]. While vision-based systems can offer powerful object detection and are already applied in famous sports like cricket [7] or tennis [8], they often require high processing power, more complex network architecture and expensive hardware. Given the need for more cost-effective and accurate alternatives, this thesis explores a different approach: testing an acoustic localization method rather than relying on optical tracking.

Acoustic Sound Source Location (SSL) is widely used in various technical applications, including automatic systems such as robots [9], drones or self-driving vehicles [10]. It finds its use in safety applications such as indoor surveillance identification [11], but also in outdoor areas to detect illegal deforestation [12] or locate a gunshot [13]. The locating process is typically realized by at least two sound-receiving sensors and a processing unit. Based on the captured audio signals and specific signal processing algorithms such as Angle of Arrival (AOA) or the Time Difference of Arrival (TDOA), the position of a sound source can be estimated in either a two- or three-dimensional space [14]. Implementing this concept, the thesis aims to build a wall positioned along the goal line of the ice hockey net. When the puck strikes the wall, the resulting sound is captured by two microphones mounted behind that wall. Existing localization algorithms are adapted to this specific puck shot scenario. As an outcome the puck's impact coordinates are determined along both the x- and y-axes of the wall. In this work, a prototype system, including both hardware and software components, is developed to implement, test, and evaluate an acoustic method for determining the location of an ice hockey puck at the moment of wall impact.

1.3. Sound Source Localization Techniques and Algorithms

Sound Source Localization is the process of identifying the origin of a specific sound signal, such as clapping, voices, or a sudden acoustic event like a hit on a wall. The determination is based on data received from microphones mounted in a known configuration, a so called microphone array. The technology is increasingly applied in areas such as robotics, autonomous vehicles, drones or home care systems, particularly where traditional systems like GPS struggle due to signal loss or reduced accuracy in obstructed indoor environments [15]. The core idea from the major global positioning systems is the same. A certain amount of receivers determine, with help of either direction-based or distance-based information, the actual location of the signal sending source. One of the most used techniques in SSL is the TDOA method. The core idea is to compute the arrival time difference between the receivers by using cross correlation, which identifies the time delay between two microphones by detecting the point of maximum similarity of their audio signals. Relating that to the known distance between the microphones and speed of sound, describes the estimated source position [14]. TDOA is advantageous compared to other methods like Time of Arrival (TOA), due to its lower computational complexity and greater flexibility for precise synchronization between microphones. Unlike TOA, which requires precise knowledge of the transmission time and tight synchronization between source and sensors, TDOA can function more flexibly by focusing on the relative time differences between microphones. Energy-based localization, in contrast, does not require synchronization at all and is commonly used in wireless acoustic sensor networks, but it lacks the spatial accuracy of TDOA [16]. Another SSL technique are DOA (Direction of Arrival) methods, which provide directional information and can operate without synchronization, such as AOA. It determines the location of the sound source using the intersection of the estimated directional ray of each sensor. In contrast to the TDOA method, which estimates the location by hyperbolic intersections [17]. TDOA presents a balanced trade-off between synchronization demands, hardware complexity, and localization accuracy, making it suitable for real-time and resource-limited applications [16]. However a single use of either method never reaches equal accuracy than the combination of both. A hybrid approach of TDOA and AOA methodologies can reduce the number of receivers in the system. Wherefore the single use of TDOA works efficiently by integrating more than four receivers, the hybrid TDOA and AOA reduces that requirement to only two receivers [17]. Because both methods result in non-linear constraints, such as trigonometric and hyperbolic constraints, solving systems like the least square method is required [18]. The prototype of the developed system of this work is provided by three microphones, acting as the receivers and also benefits from the additional known constrain of the wall's position.

2. Materials and Methodologies

2.1. Hardware Setup

2.1.1. Construction and Sensor Integration at the Goal Wall

A simple and stable fixation of the wall along the goal line is essential for the system's ease of use. The mounting process must be quick and straightforward. For this purpose, a clamping system secures a rectangular wooden frame to the goalposts and crossbar of a standard ice hockey goal (Figure 2.2c). In order to simplify the fabrication process and obtain the required material characteristics rapidly and functionally for the prototype, the clamps are manufactured additively with a *Prusa MK4S* 3D-Printer. They are made of Polylactide (PLA) and specifically designed to provide a secure and stable grip on the goal framework. Due to its ability to deliver high mechanical performance when printed with optimized settings, PLA is highly suitable for rapid prototyping. Using a 0.2 mm layer height, a printing speed of 20 mm s^{-1} , and 1.75 mm filament, the print achieved good inter layer adhesion and structural stability, making it well-suited for applications requiring rigidity and flexure strength [19].

After designing the parts in the CAD program (Computer Aided Design) *Autodesk Inventor*, the file is exported as an STL (Standard Triangle Language) file, and with the Prusa Slicing Software, converted to a for the printer readable G-Code. While the first prints were lacking on stability, the improved version is now designed with a slightly bigger diameter on the outside (55 mm) and a decreased closed angle of 100° . Figure 2.1 gives more detailed pictures of the design and dimensions. As a result, the clamps fit tight around the standardized posts and crossbar and offer a simple mounting opportunity for every user.

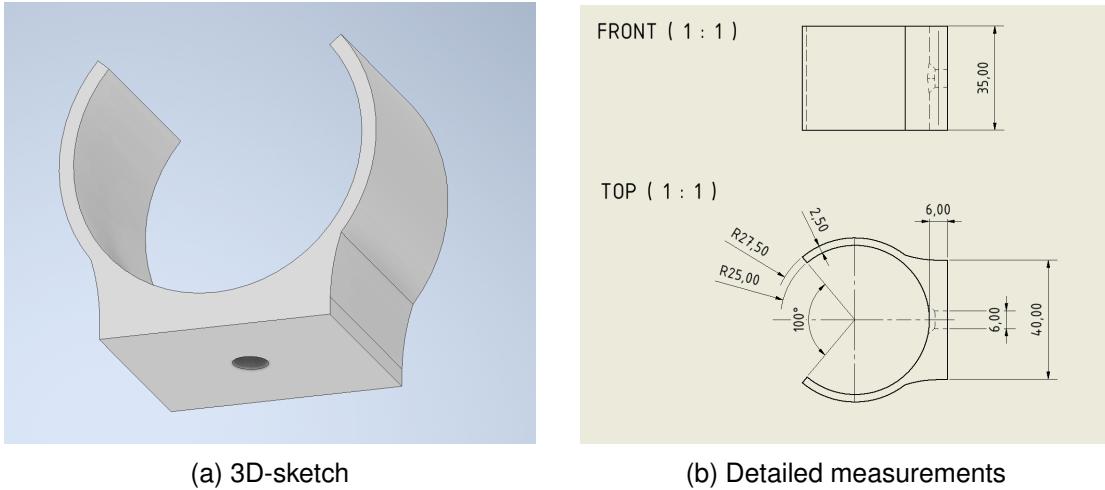


Figure 2.1.: Clamp design and measurements: a) shows a 3D image of the Design, while b) is giving insights about the exact dimensions of the clamp.

Since the dimensions of the ice hockey goals are standardized, the outer frame is 190 cm wide and 126.5 cm high and covers the whole goal, including its framework. The wooden frame offers the surface for a simple wall installation, while the screwed on clamps lock it firmly in place by snapping around the rounded posts and crossbar. Instead of fastening the wooden wall directly onto the frame, a thin layer of foam is glued between. This layer is intended to dampen high-impact noise and support signal processing by absorbing part of the vibrational energy, particularly in the high-frequency range. The wooden wall, which matches the frame dimensions, is made from pine and has a thickness of 10 mm, thereby combining good structural rigidity, while lowering damping effects due to its minor mass. It is anchored firmly to the frame using twenty symmetrically distributed wood screws. Figure 2.2c gives a detailed idea of all the layers mounted.

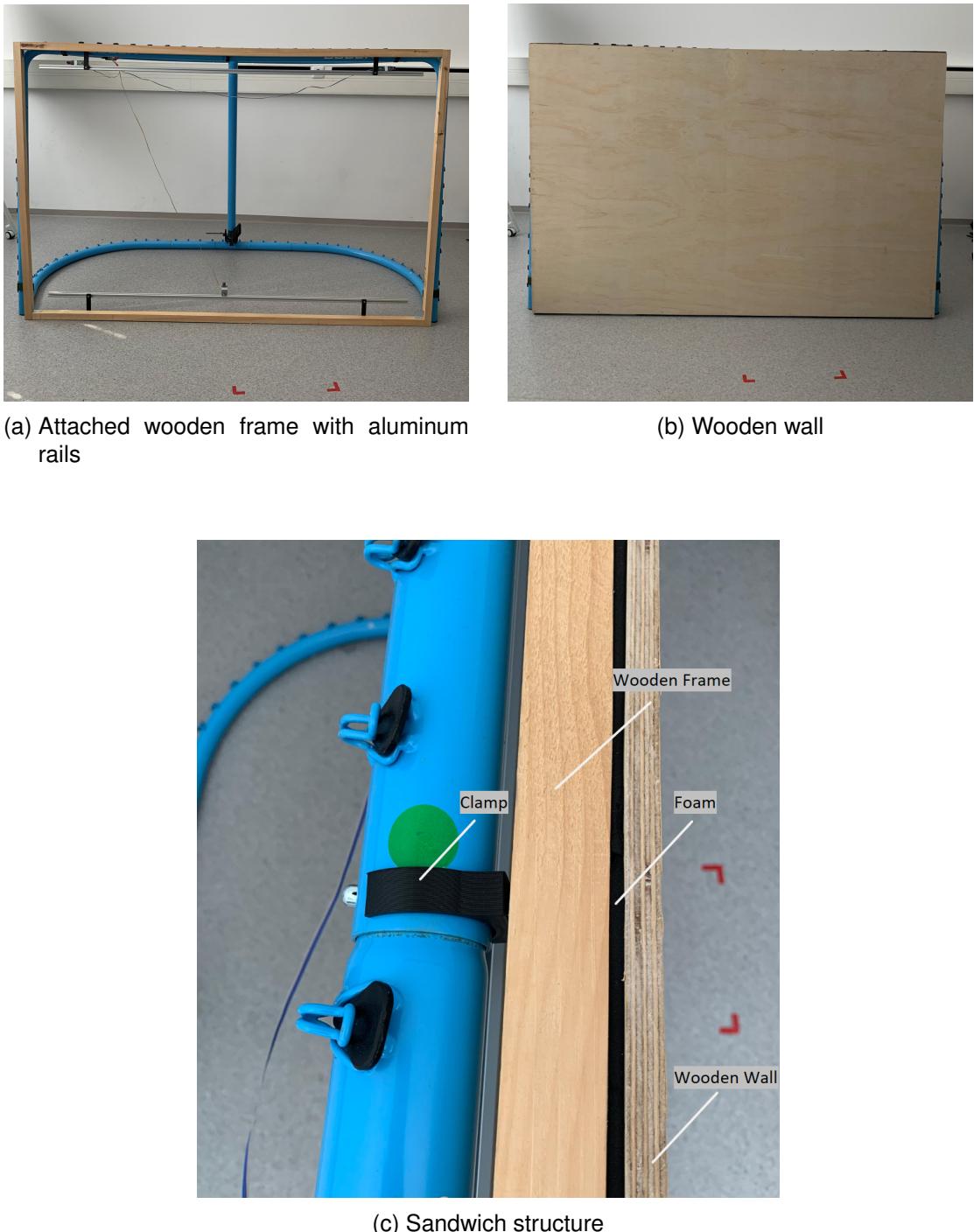


Figure 2.2.: Frame and wall mounting setup: a) Attached wooden frame with aluminum rails, onto which the microphones are mounted, b) wooden wall attached to the frame, and c) sandwich structure from the top view.

The wooden frame not only keeps the wall in place, it also offers the base to attach two horizontally and parallel aligned aluminum profiles, which function as a flexible mounting slide for the microphones (Figure: 2.2a). Four 3D-printed rectangular parts position the rails 8 cm away from the bottom or top. All parts are made from the same material (PLA) and were also printed using the *Prusa MK4S* 3D-Printer. With a thickness of

6 mm, they offer sufficient flexure strength and stability to hold onto the light weight of the aluminum profiles. The profiles are measured with the standard size of 40 mm width and height and reach a length of 170 mm. As a result, the microphone adapter can be securely positioned and easily stabilized at any point along the rail. Just by tighten the M4 Countersunk Head Screw with an Allen key, the microphone sits accurate at the desired position.

2.1.2. Microphone Framework

Microphones are widely used across various applications, transforming sound waves into electrical signals. They are classified based on their technology, including piezoelectric, piezo resistive, optical, and capacitive mechanisms. The key difference lies in their transducer types, directionality and frequency response. The piezoelectric microphone works based on the piezoelectric effect. Sound waves apply stress and cause charges on piezoelectric plates, which can be measured as the output voltage. The piezo resistive function instead relies on the specific semiconductor material which changes its electrical resistance due to the sound waves. A different approach to realize transduction is an optical one. It works by detecting changes in the light intensity and is mostly used when noise canceling is needed. The study uses capacitive microphones because they are valued for their high sensitivity and flat frequency response, making them suitable for this project. They divide into two types of microphones: a simple condenser microphone and the electret condenser microphone. The term electret refers to a material that maintains permanent electrical charge on the built-in plates [20].

Microphones capture sound differently based on their directional characteristics, which determine their sensitivity to various angles. These characteristics are represented through polar pattern plots, illustrating how sound is received from different directions. Omnidirectional microphones pick up sound equally from all directions, ensuring the same sensitivity to the sound. This makes them particularly useful in situations where the sound source changes location. In contrast unidirectional microphones are most sensitive to sound from a single direction following only one axis [21].

The chosen MAX9814 microphone is an integrated microphone amplifier designed for electret condenser microphones. It features a low-noise preamplifier, automatic gain control (AGC), and an internal 2 V bias voltage generator to power the microphone. The AGC with adjustable output gain ensures consistent output levels by dynamically adapting the gain to the input signal strength, reducing distortion and preventing clipping. The board contains four hardware pins and operates with a voltage from 2.7 V to 5 V. The *Vdd* and *GND* pins are soldered to their corresponding ports of the microcontroller. The *Out* realizes the analog data connection to the Teensy. *AR* describes the Attack and Release Time Ratio for the AGC circuit and determines how rapidly the AGC responds to increases (attack) and decreases (release) at the input signal level. The *Gain* sets the amplifier magnitude for the output signal. Like the *AR* pin, can the

Gain pin be set to three states (Figure: 2.1) [22]. Both remain in float for the first tests, but can be will be adapted later to improve signal quality. Since the impact events can be very loud and result in magnitudes that exceed 3.3 V, the AGC of the microcontroller reaches its limitation. Due to that clipping effects in the raw signal can appear in form of plateaus at the maximum values. Reducing the *Gain* can prevent this effect and should be considered in signal processing.

Table 2.1.: Pin Setting for Gain and Attack/Release Time Ratio.

| Connection | Gain Setting | A/R Time Ratio |
|------------|--------------|----------------|
| GND | 50 dB | 1:500 |
| Vdd | 40 dB | 1:2000 |
| Float | 60 dB | 1:4000 |

In order to make the MAX9814 board more flexible, it is soldered on a regular circuit board. All soldered connections are protected with heat shrink tubing and tape serves as strain relieve for the contacts. The wires have a diameter of 0.64 mm, ensuring the signal's integrity due to the low resistance and moderate length under 2 m. A L-shaped corner bracket, screwed to the circuit board, completes the adapter and functions as an easy fixation (see Figure 2.3).



Figure 2.3.: Microphone attachment views: a) Back view and (b) front view of a top microphone mounted on the aluminum rail.

As a result, all three microphones attach to the aluminum profiles. The upper ones are placed in directing the corner and are horizontally aligned, while the third one is mounted centrally on the lower profile. In general, two or more microphones form either a linear or planar microphone array and serve as the receiving system to detect and determine the direction of incoming sound [15]. In this case, the microphones realize a planar and triangular configuration. Unlike a linear setup, which limits DOA estimation to a single dimension, a triangular array provides spatial diversity in two dimensions

and thereby enhances directional sensitivity, allowing more accurate localization on a plane [23].

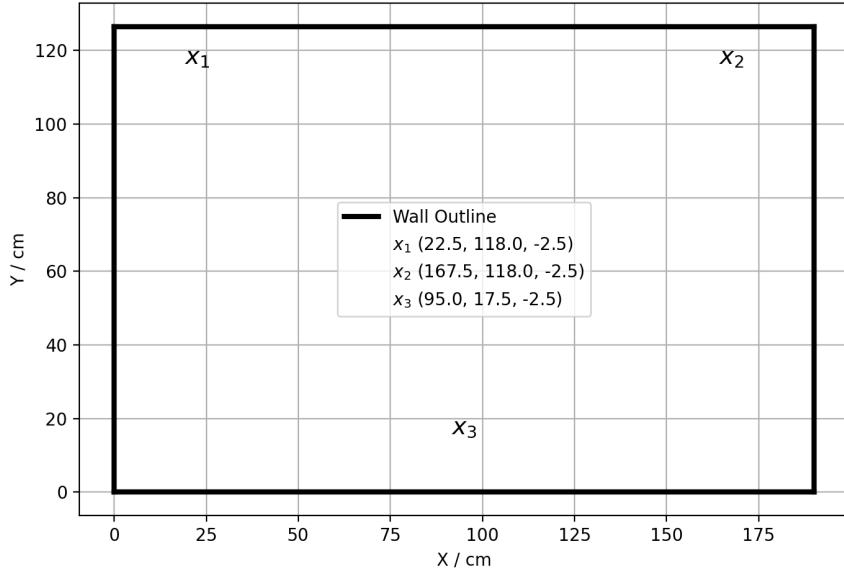


Figure 2.4.: Planar and triangular microphone configuration from a two dimensional point of view

2.1.3. Microcontroller and Microphone Communication

The microcontroller builds the crucial interface between the software logic and the hardware components. It processes the audio data of the three microphones and provides the SSL Python script with the recorded impact data. The important characteristics of this project are the sampling frequency and hardware limitations for communication with the microphones. Because of its fitting requirements, the ESP32 is chosen for a first try out. This low-cost and energy-efficient device serves as the core component in various specialized boards designed for tasks like real-time audio processing [24]. Regarding the project that achieves SSL on the ice hockey goal wall, the ESP32 seems suitable due to its sampling frequency of 44.1 kHz and the I2S (Inter-IC Sound Bus) communication protocol. The protocol was created by Philips Semiconductors to simplify the integration of digital audio components and support robust design implementations. It is a serial communication protocol designed for audio data. Two stereo channels operating synchronously using a single data line for the audio streams.[25] Tests of the ESP32 with the I2S setup in combination with three INMP441 MEMS Microphones show lacks in its functionality. While running two microphone inputs simultaneously works out well, problems occur when adding a third data line to the system. Despite testing various software approaches, the solution for running three synchronized independent audio streams from three separate microphones does not function as intended.

Due to this hardware limitation, the system gets revised. The prototype's next generation (Stage 1) reads out three analog MAX9814 microphone inputs with dual ADC

(Analog to Digital Converter) sampling. Since the ESP32 only covers a single ADC, the microcontroller is replaced with a Teensy 4.1, which provides two independent 12-bit ADC's allowing near simultaneous sampling, additional 14 analog inputs and a power supply of 3.3 V to 5 V for the microphones. The audio data are captured and stored in a CSV-file (Comma-separated-values) with one column for each microphone output. A serial connection builds the bridge between the calculation software and the CSV-Datafile. The output of each microphone is soldered to one of the first three analog inputs of the Teensy 4.1 (A0, A1, A2), along with connections for the 3 V power supply and ground, as shown in Figure 2.5.

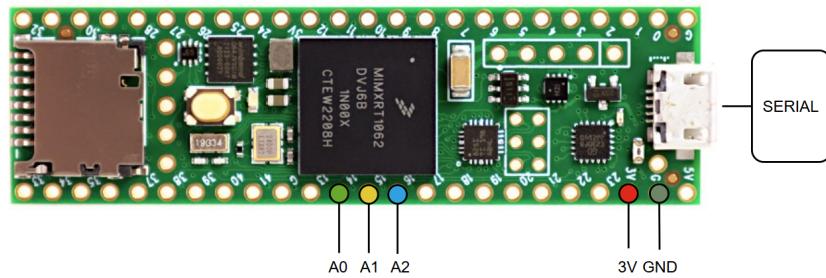
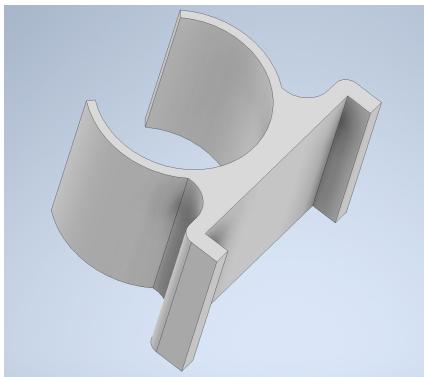
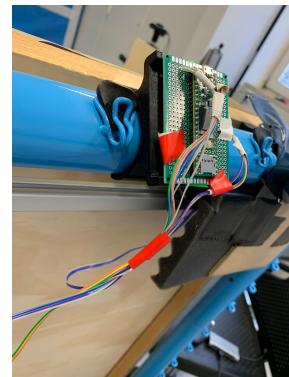


Figure 2.5.: Used ports of the Teensy 4.1

The microcontroller itself is soldered on a circuit board and attached to an adapter part, that clamps both to the goal behind the wall to ensure its protection (see Figure 2.6). The adapter can easily be clipped to any aluminum part of the goal, such as the crossbar or the round bars behind the goal wall. It is fabricated in the way as the other printed parts mentioned before.



(a) 3D sketch



(b) Crossbar attachment

Figure 2.6.: Teensy adapter and mounted configuration: a) shows the Design of the adapter clamp, and b) visualizes how the Teensy is attached and mounted to the hardware.

2.2. Software Setup

2.2.1. Code Architecture

The code architecture divides into three main parts. Firstly, the logic of the Teensy 4.1 to read out the three analog data inputs of the microphones and synchronize this information. This is important to determine the TDOA of the signal and gain precise time relations between every possible microphone pair by applying cross correlation. Secondly, a script processing the serial sent data, storing and labeling it correctly in a CSV-file. And as the crucial part of the code, a third script containing the signal processing, filtering, followed by the calculation and SSL by applying the hybrid approach of combining TDOA and AOA methodologies. The result of the impact location is visualized graphically.

2.2.2. Capturing Audio Data

The programming language of the written code for the Teensy 4.1 is C++. Therefore, *PlatformIO IDE* is installed in the *VisualStudio* development platform. This extension facilitates building the code in the required Arduino framework. The script realizes real time audio signal recording and stores the data buffer after an impact event in an CSV file. Therefore it sets constant values like the sampling frequency, buffer size and configures the pins to the analog inputs A0, A1 and A2. The hardware initialization enables the dual ADC use for parallel reading the inputs A0 and A1 of two microphones. The third microphone is read out sequentially due to hardware limitations. As a result the audio data of the third microphone is read as well at ADC0 after the first microphone. This causes an delay of a few microseconds. The conversion time depends on the combination out of the resolution (12 bit), the ADC's clock speed and the conversion speed setting, which is implemented in the code as the maximum. In the project's use-case the conversion time for a 12 bit resolution and *VERY_HIGH_SPEED* is typically measured between 4 μs and 6 μs [26].

$$\Delta n_{\text{mic}3} = \Delta t_{\text{mic}} \cdot f_s = 5 \mu\text{s} \cdot 44\,100 \text{ Hz} \approx 0.22 \text{ samples} \quad (2.1)$$

Equation 2.1 shows the resulting delay in samples when using a sampling frequency f_s of 44 100 Hz. In relation to the speed of sound in the air, the actual spatial offset can be determined (see Equation 2.2). Table 2.2 provides an overview of the small spatial shifts resulting from the delay. The minimal shifts do not require further software correction.

$$\Delta d = c \cdot \Delta t_{\text{mic}} = 0.343 \frac{\text{mm}}{\mu\text{s}} \cdot 5 \mu\text{s} = 1.715 \text{ mm} \quad (2.2)$$

Table 2.2.: Conversion of time delays to sample delays and equivalent spatial errors

| Time Delay / μs | Sample Delay | Spatial Error / mm |
|----------------------------|--------------|--------------------|
| 5 | 0.22 | 1.72 |
| 10 | 0.44 | 3.43 |
| 15 | 0.66 | 5.15 |

Additionaly constants like the *PRE_IMPACT* and *POST_IMPACT* define how much data is captured before and after an impact, ensuring the full event is recorded. To isolate each impact and avoid recording additional echos or artifacts a *COOLDOWN_MS* delay prevents the system from reacting to rebound signals.

To save the data from the received buffers, the python script *save_captured_impacts* builds the serial connection to the microcontroller using *pyserial*. Once an Impact is detected, it collects the three-channel data from the microphones which cover 512 samples each and saves it as a CSV-file in the desired folder.

2.2.3. Signal Filtering and Processing

The CSV-file with the three raw audio data channels needs to be further processed to determine the sound location on the wall. Each impact on the wall passes each step of the process, visualized in Figure 2.7.

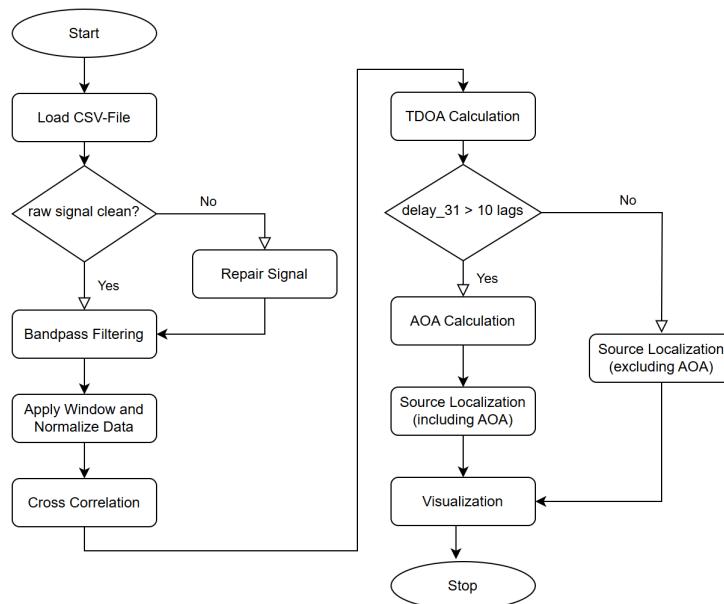


Figure 2.7.: Flowchart Signal Processing

The significant impact of the signal processing is emphasized by comparing the plotted raw signal against the processed one before the TDOA Estimation (Figures 2.8 and 2.11). For accurate cross correlation and reliable TDOA estimation, the raw signal must

be free of clipping and clearly capture the critical segments of the signal peaks. Clipping describes the distortion that arises when the amplitude of an audio signal exceeds the maximum level that a microphone or ADC can handle [27]. The first measurements include typical clipping plateaus at the maximum and minimum values of the microcontroller's ADC. To reduce this effect the microphones default *Gain* setting get changed from 60 dB to 40 dB. The outcome does still not provide a clipping free signal, especially when sound events are louder such as a shot with the ice hockey puck compared with a knock at the wall (see Figure 2.8). To address the issue of clipped signal segments, interpolation methods such as spline-based reconstruction can be applied within the software to estimate and restore the missing waveform portions caused by amplitude saturation [27]. Figure 2.9 shows the result of the approach of reconstructing clipped signal segments. After normalizing and rescaling the raw signal, three implemented functions in the script work together to detect the clipped samples of the signal, group them into continuous ranges, and smoothly reconstruct clipped regions by using cubic spline interpolation.

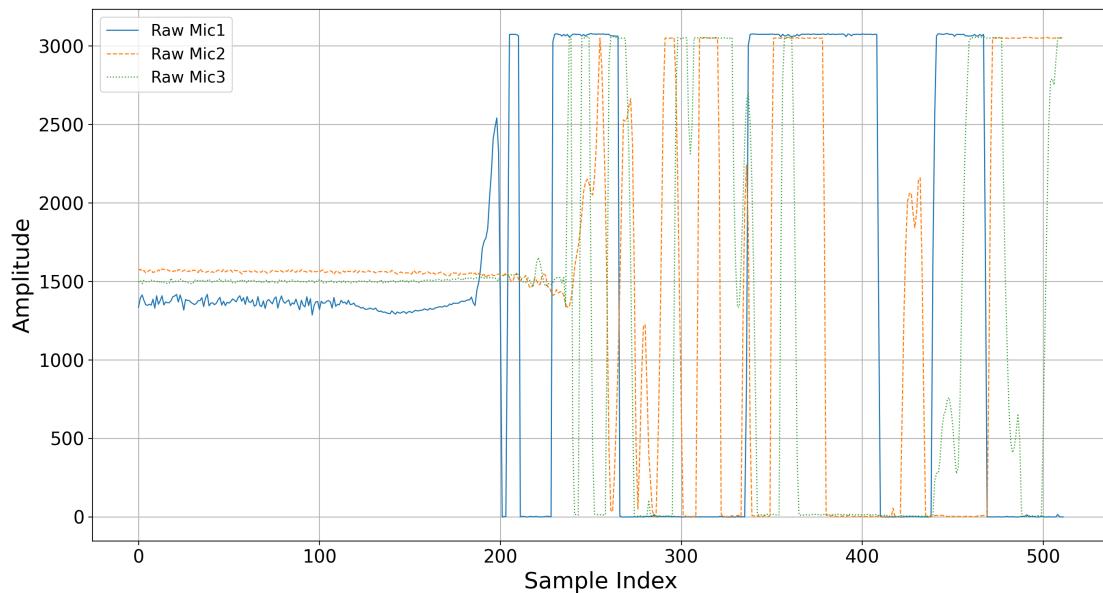


Figure 2.8.: Raw microphone signals

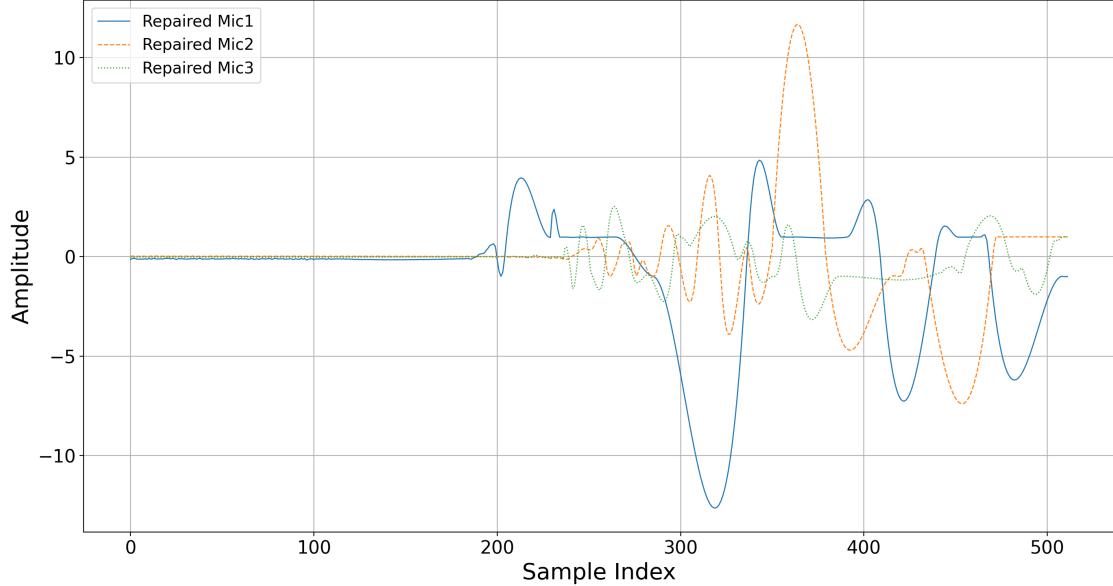


Figure 2.9.: Repaired microphone signals

The further signal processing applies bandpass-filtering and ensures that relevant frequencies are included in the further calculation. The goal is to exclude background noise or ringing effects from potential reflection or echos. To set appropriate cut off frequencies for the bandpass filter, the frequency band of the puck's impact where the majority of energy is concentrated, is the key insight. Welch's method is applied to estimate the power spectral density (PSD) of recorded impacts. This method offers a smoothed and statistically stable view of the signal's frequency content by averaging modified periodograms of overlapping segments [28]. By analyzing the distribution, the dominant components of the signal can be identified and cutoff frequencies that preserve the essential information while attenuating irrelevant low-frequency drifts and high-frequency noise can be selected.

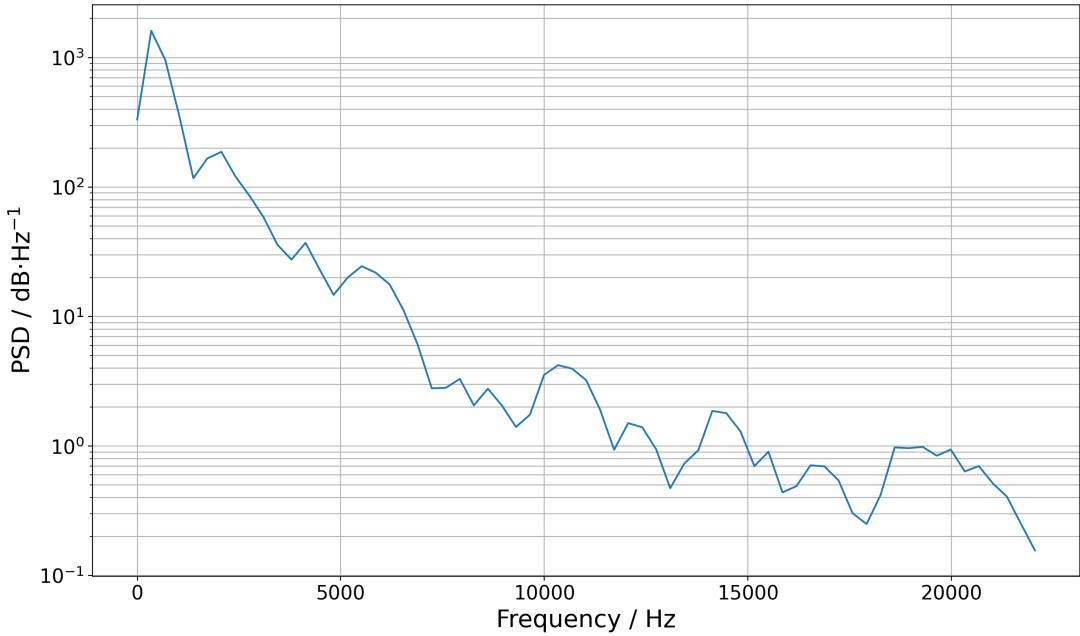


Figure 2.10.: Welch's method: Puck's energy distribution

Figure 2.10 provides insight into the dominating frequencies where most of the energy of the shot is to be found on the wooden wall. The dominant energy lies between 1 kHz and the range of 4.5 kHz to 7 kHz. This suggests applying cut-off frequencies from 0.8 kHz and 6 kHz.

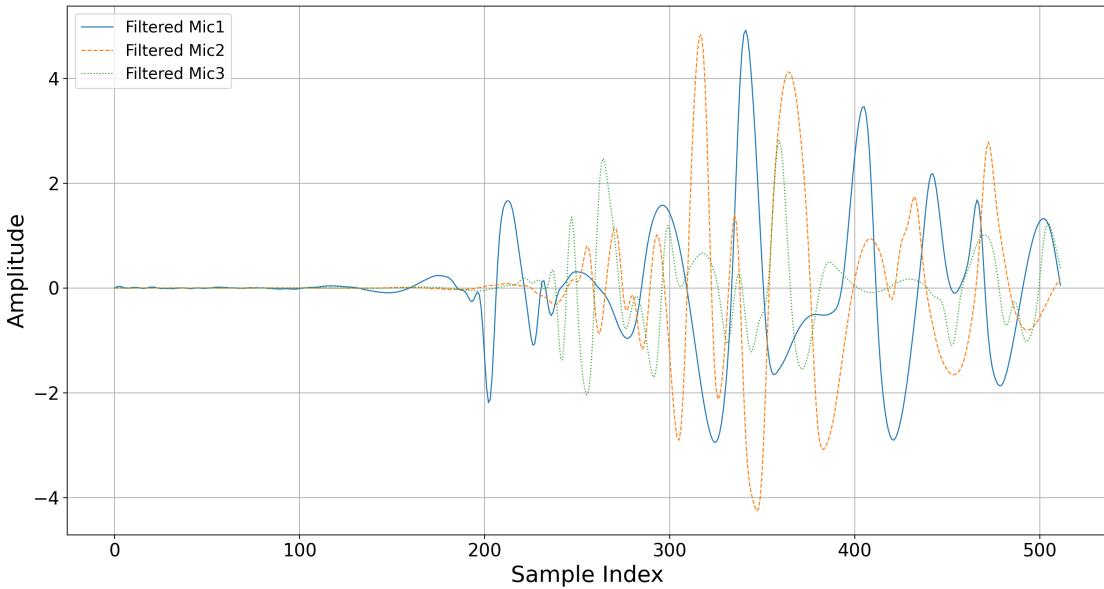


Figure 2.11.: Bandpass filtered signal

Additionally, a window cuts out irrelevant parts of the signal since only the main peaks, which are seen as the actual impact peak, are crucial for the delay estimation. This reduces the risk of false detections in the cross correlation. This step is only done at

the reference microphone 1 (Figure 2.12).

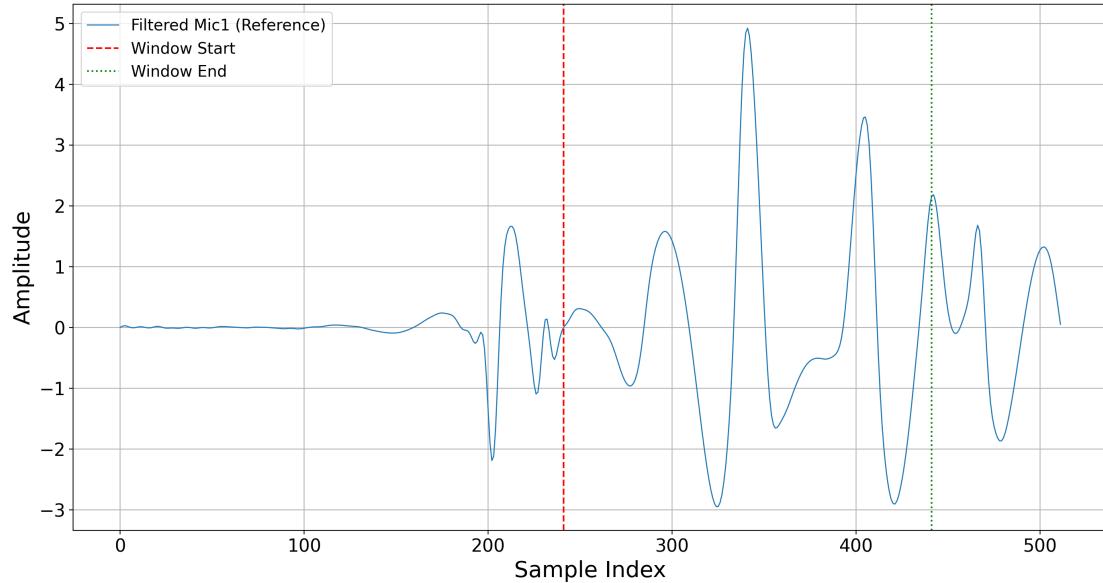


Figure 2.12.: Windowed signal with plus and minus 100 samples around the detected peak

2.2.4. Calculation Methodology and Location Determination

The TDOA estimation is based on the delay in lags of when the sound arrives at the microphones. With this value either the time difference of arrival or the delay in samples can be calculated. To determine this specific delay, the correlate function of *SciPy* library is used. With help of the Fast Fourier Transformation (FFT) method, the output is an one-dimensional array which contains the best alignment of the two signals. The peak of this correlation function corresponds to the point of strongest similarity, caused by the sound impact. The lag at which this peak occurs represents the sample delay between the signals and is used to compute the TDOA [29]. Therefore, the windowed signal of reference microphone 1 is shifted along all the filtered samples of the other microphones, summing the result at each lag. Where the peak is the highest, the alignment is the most similar and the delay in lags can be determined. Figure 2.13 and Figure 2.14 visualize the highest peak after applying the correlate function to the signal pairs. One lag is referred to $22.68 \mu s$ (see Equation 2.3).

$$1 \text{ lag} = \frac{1}{f_s} = \frac{1}{44\,100 \text{ Hz}} \approx 22.68 \mu s \quad (2.3)$$

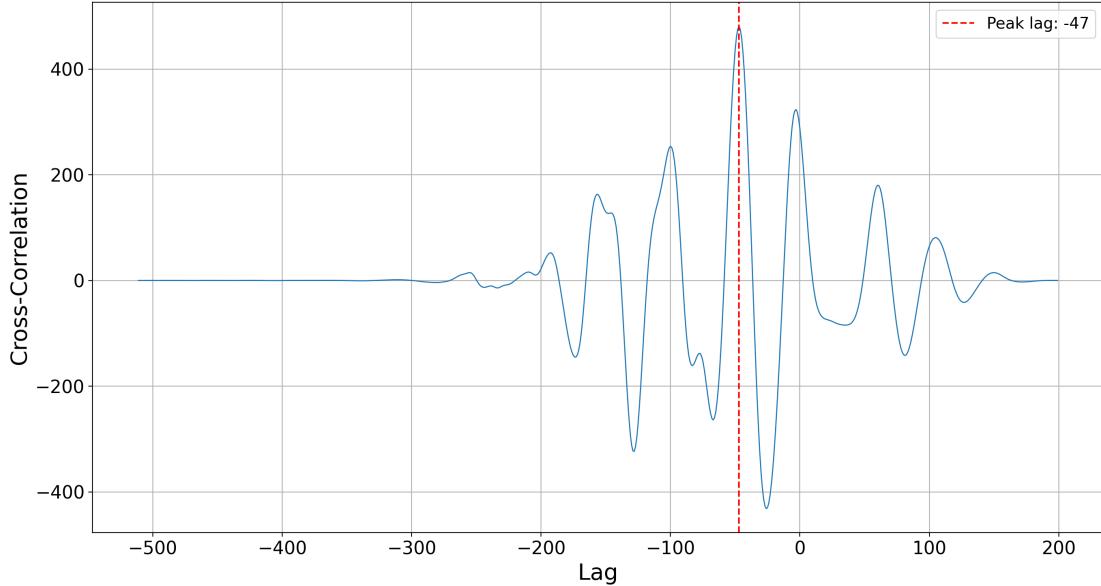


Figure 2.13.: Cross Correlation between microphone 1 and 2

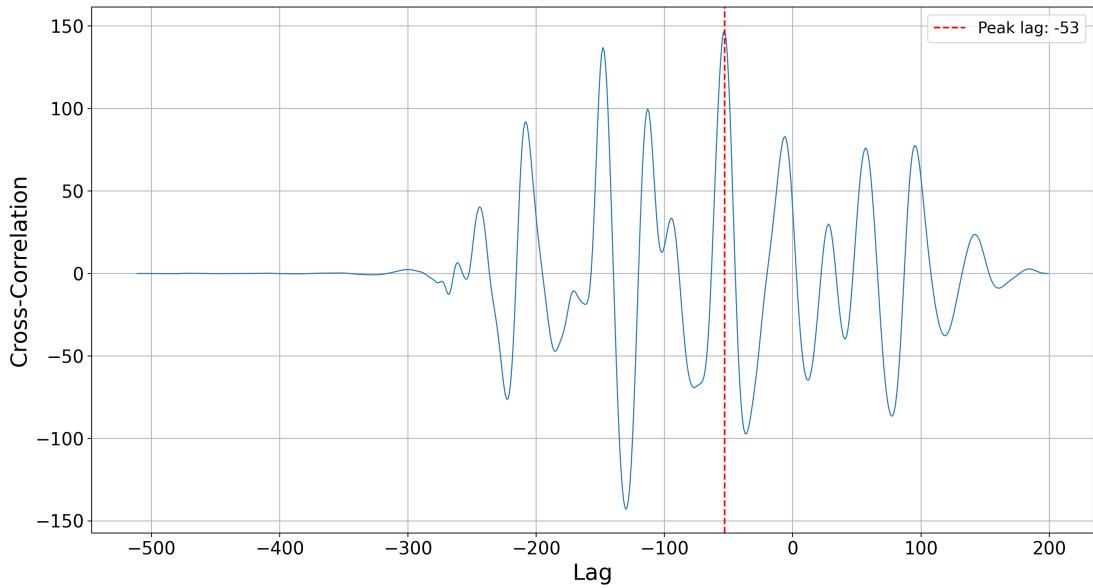


Figure 2.14.: Cross Correlation between microphone 1 and 3

The calculation of TDOA is represented in Equation 2.4. The values r_{21} and r_{31} represent the ranges of distances in cm , that the sound waves travel from the source between microphone 1 and microphone 2 or 3. Assuming sound as a constant, it gets multiplied with the TDOA value, shown in Equation 2.5. Microphone 1 is chosen as the reference, and only the time differences relative to this microphone are required for SSL. The additional constraint that the wall lies in the x/y-plane ($z = 0$) reduces the solution space from three to two dimensions.

$$tdoa_{21} = \frac{delay_{21}}{f_s}, \quad tdoa_{31} = \frac{delay_{31}}{f_s} \quad (2.4)$$

$$r_{21} = c \cdot tdoa_{21}, \quad r_{31} = c \cdot tdoa_{31} \quad (2.5)$$

The AOA method builds the hybrid with the TDOA and adds an extra constraint to the system with the goal of increasing its accuracy. The $toda_{31}$ can also be used to calculate the AOA of the sound wave using Equation 2.6 where D is the distance between the microphones, known as nodes, and c is the speed of sound (343 m/s) [14].

$$\theta_{31} = \arcsin \left(\frac{toda_{31} \cdot c}{D} \right) \quad (2.6)$$

All the calculated and given constraints realize the arguments for a cost function which is used to solve the non numerical equation by using the least squares method to estimate the SSL correctly. Since the positions of the microphones (m_1, m_2, m_3) are known, the geometric distances from the source to each microphone are defined as:

$$d_i = |x - m_i|, \quad i \in 1, 2, 3 \quad (2.7)$$

Given the measured range differences from the source to the corresponding microphone pair (r_{21}, r_{31}) the first two terms of the cost function.

$$term_1 = (d_2 - d_1 - r_{21})^2 \quad term_2 = (d_3 - d_1 - r_{31})^2 \quad (2.8)$$

The AOA estimation adds an additional third term to the function. σ describes weight of the AOA estimation in the equation. If the delay in lags is lower than 10, the reliability for the estimation decreases and σ down weights $term_3$ in the cost function (see Equation 2.9).

$$f(x) = term_1 + term_2 + term_3 = (d_2 - d_1 - r_{21})^2 + (d_3 - d_1 - r_{31})^2 + \left(\frac{\theta_{\text{meas}} - \theta_{\text{est}}}{\sigma} \right)^2 \quad (2.9)$$

The cost function runs an optimization algorithm that tests many possible impact locations and compares the predicted delays to the measured ones. The sound source location is found by minimizing the total squared error across all microphone pairs [30].

The difference in distances between two microphones defines a geometric constraint in the form of a hyperbola. All possible source locations that correspond to a specific TDOA lie along this curve. The cost function minimizes the squared difference between the measured and theoretical distance differences and finds the position fitting the constraint best as possible. Figure 2.15 visualizes this fitting process with a heatmap, where the best possible fit is represented in the darkest area of the map.

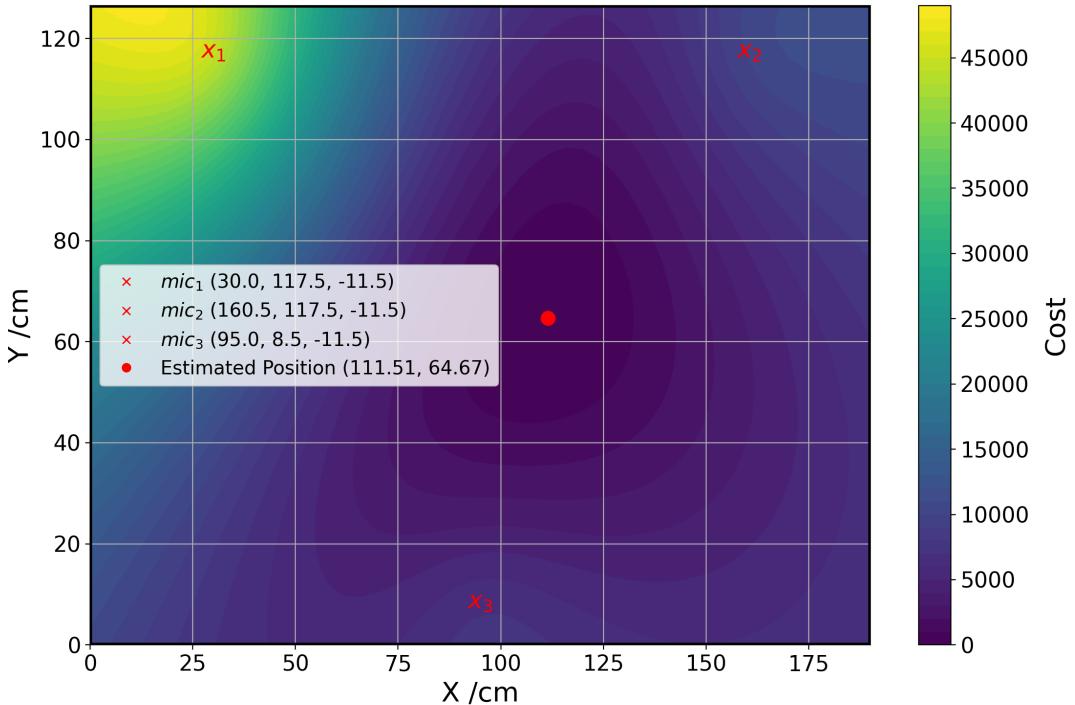


Figure 2.15.: Heatmap of the cost function to estimate the position

2.3. System Testing and Data Recording

To evaluate the system's accuracy, two test series are conducted, which are applied to different states of the prototype. In the first series, nine predefined positions (see Table 2.3) on the wooden wall are marked (see Figure 2.16a), and an ice hockey puck is manually thrown at each of these targets. The systems origin lies in the bottom left corner. This allows for a systematic validation of the position estimation accuracy across the entire wall surface, as each impact point has a known reference location.

Table 2.3.: Coordinates of nine marked positions

| Position | (x, y) |
|----------|------------|
| P1 | (30, 110) |
| P2 | (90, 110) |
| P3 | (150, 110) |
| P4 | (30, 80) |
| P5 | (90, 80) |
| P6 | (150, 80) |
| P7 | (30, 40) |
| P8 | (90, 40) |
| P9 | (150, 40) |

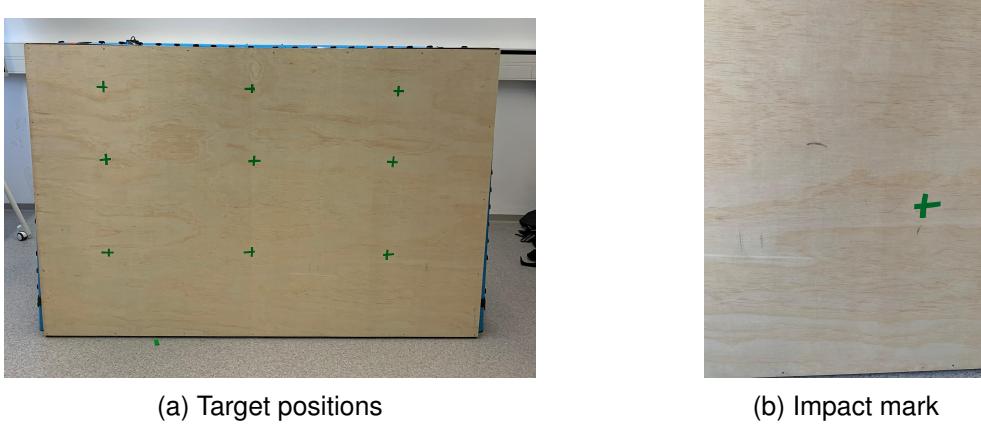


Figure 2.16.: Reference points on the wall and shot impact mark: a) shows the nine green marker positions used as reference targets for the throwing series, and b) visualizes the black puck mark, which is used to measure the real coordinates of the hit after a shot.

In the second series, multiple shots are taken using an ice hockey stick under realistic conditions. The puck leaves a visible mark on the wall after an impact (see Figure 2.16b), enabling the true coordinates of the hit to be measured and compared to the positions estimated by the system.

Additionally, the delay in lags is recorded for each impact event. As the core principle of the system relies on TDOA estimation, these delays play a crucial role in accurately determining the source coordinates. Comparing the measured delays with the theoretically expected lag values provides an effective metric to evaluate the accuracy of the localization algorithm. Another metric for a better analysis is the AOA and is also calculated theoretically, depending on the referent impact location.

First, the distances from the source to two microphones are calculated, where s is the position vector of the sound impact location, and m_i , m_j are the positions of a microphone pair.

$$d_i = |s - m_i|, \quad d_j = |s - m_j| \quad (2.10)$$

The time delay between this microphone pair is calculated and then converted into the lag delay using the system's sampling frequency f_s .

$$tdoa_{ij} = \frac{d_j - d_i}{c} \quad (2.11)$$

$$lag = tdoa_{ij} \cdot f_s \quad (2.12)$$

This theoretical lag can be directly compared with the measured lag obtained from cross correlation to validate the accuracy of the TDOA-based localization.

The theoretical is calculated with the vector between m_1 and m_3 . The angle is then obtained using the arctangent function:

$$\theta_{\text{AOA}} = \arctan 2(y_s - y_1, x_s - x_1) \quad (2.13)$$

where x_s and y_s are the coordinates of the sound source and x_1, y_1 is the position of microphone 1. This angle describes the absolute direction of the source with respect to the horizontal axis, measured in radians. It is converted to degrees for better interpretability with Equation 2.14.

$$\theta_{\text{deg}} = \theta_{\text{AOA}} \cdot \left(\frac{180}{\pi} \right) \quad (2.14)$$

The first round of testing series is applied to the prototype's state described above (Stage 1). A second testing round (Stage 2) is conducted to evaluate the effect of potential improvements, such as shielding the microphones from reflections and echoes using acoustic panels (see Figure 2.17). These panels allow direct sound paths within the setup while blocking external reflections to reduce strong echo effects from the sides or back.



(a) Side view of microphone 2



(b) Top view of microphone 3

Figure 2.17.: Microphones shielded by acoustic panels: a) shows the side view of microphone 2, shielded from the outer sides and the back and figure b) represents the top view of microphone 3, which is shielded from the back and ground.

An additional software update enhances the signal processing by applying the windowing step not only to the reference microphone but to all microphone signals (Stage

3). Cross correlation is then computed only within this shared impact detecting window, which improves alignment and robustness. Moreover, the update introduces constraints on the maximum allowable sample range during delay estimation, based on the microphone distances of the setup (see Equation 2.17). These settings realize the prototype in stage 3.

$$d_{ij} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (2.15)$$

$$d_{12} = \sqrt{(167.5 - 22.5)^2 + (118 - 118)^2 + (-2.5 + 2.5)^2} = \sqrt{145^2} = 145 \text{ cm} \quad (2.16)$$

$$\text{delay}_{\max} = \frac{d_{12}}{c} \cdot f_s = \frac{145 \text{ cm}}{34300 \text{ cm/s}} \cdot 44100 \text{ Hz} \approx 186 \text{ samples} \quad (2.17)$$

In a last development step (Stage 4), the microphone configuration is adapted with the goal to increase TOA and decrease the frequency of arriving echoes and ringing effects through longer distances between the wooden wall. The following Table 2.4 declares the new microphone coordinates.

Table 2.4.: Triangular microphone configuration with adapted coordinates in stage 4.

| Microphone | Position (x, y, z) |
|------------|-----------------------|
| Mic 1 | (30, 117.5, -11.5) |
| Mic 2 | (160.5, 117.5, -11.5) |
| Mic 3 | (95.0, 8.5, -11.5) |

In addition, the windowing function is improved by detecting the first significant increase and not the maximum peak of the signal. The input signal is first smoothed using a moving average to suppress noise. The derivative of the smoothed signal is then calculated to highlight steep signal rises. The algorithm searches for points where the derivative exceeds a defined threshold and the amplitude is sufficiently high. To avoid false detections from initial fluctuations, only samples after a specified index are considered. If no suitable point is found, the maximum peak of the raw signal is used. Again a fixed number of samples before and after the detected peak defines the window (see Figure 2.18).

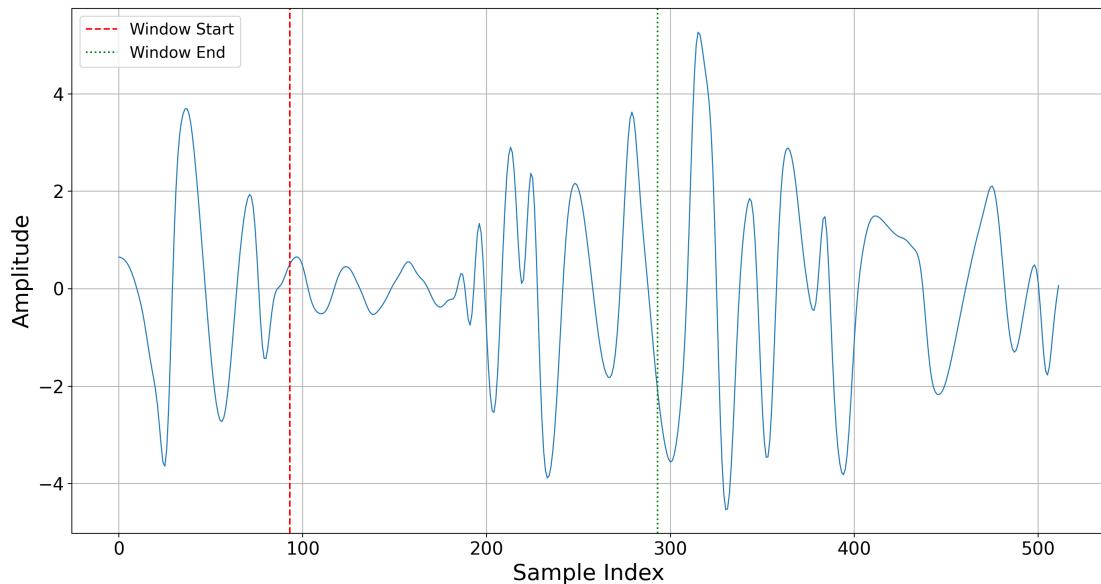


Figure 2.18.: Window detection of the most relevant peak and not the maximum at stage 4.

3. Results and Analysis

The test results are compared in their different stages of development:

- **Stage 1** - undamped; windowing at the signal of reference microphone 1 only
- **Stage 2** - damped setup, windowing at the signal of reference microphone 1 only
- **Stage 3** - damped setup, windowing applied to all microphone signals
- **Stage 4** - damped setup, adapted microphone configuration; enhanced windowing function

The localization error, calculated according to Equation 3.1, evaluates the performance of the system and the impact of iterative adaptations of hardware and software components at each stage of development. It reflects the accuracy of each prototype in centimeters.

$$\text{error} = \sqrt{(x_{\text{est}} - x_{\text{true}})^2 + (y_{\text{est}} - y_{\text{true}})^2} \quad (3.1)$$

3.1. Localization Accuracy Development

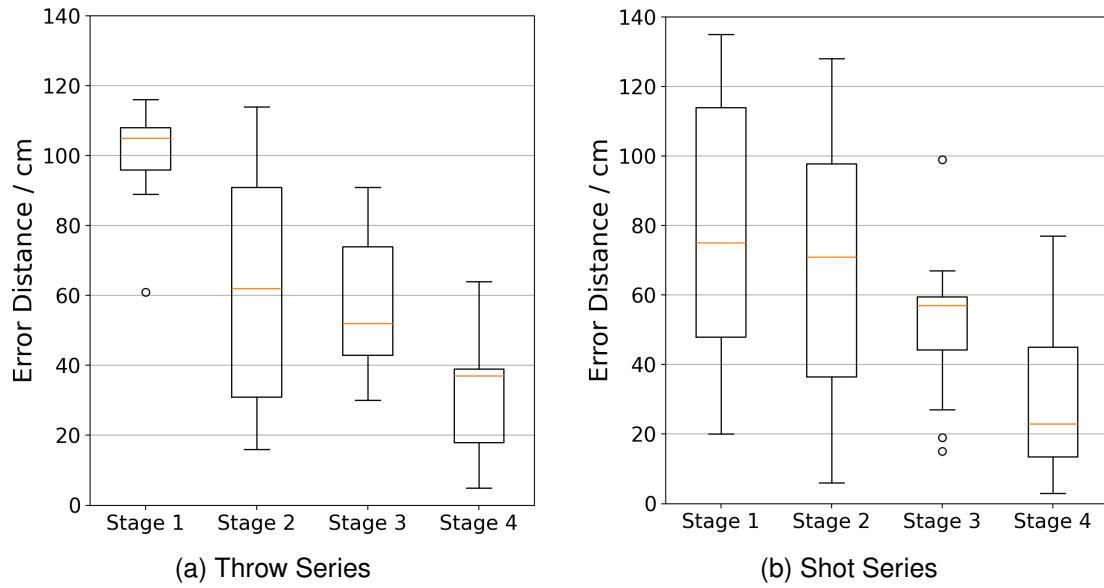


Figure 3.1.: Error distance visualization: a) illustrates the Throw Series, while b) is representing the Shot Series. Both include the development across all four prototype stages and show the metrics in Boxplots.

The Box Plots in Figure 3.1a and Figure 3.1b visualize the localization accuracy across all four development stages. They indicate a clear trend of positive performance development. In both test series, the boxes illustrate a reduction in the error magnitude and variability in each stage. The relevant metrics are given in Table 3.2 and Table 3.1 and emphasize the visual interpretation. The median error in the throw series decreases from 105cm in Stage 1 to 37cm in Stage 4, while in the shot series, it drops from 75cm to 23cm . Remarkable is that especially the damping had a huge impact for the throwing test series. From stage 1 to stage 2 the error could be reduced by 40.95%, while it was less remarkable at the shot series. Another significantly enhancement can be observed in stage 4. The combination of the new microphone configuration and the improved peak detection led to a decrease in localization error of approximately 59.65% for the shot series and 28.85% for the throwing series. This emphasizes how essential functional cross-correlation is and highlights its relevance from a fully optimized system. The improved spatial resolution led to better signal quality, and in combination with a more accurate peak detection method, delay estimation became significantly more precise. By comparing Table A.2 and Table A.4 in relation with Table A.6 and Table A.8, the improvement of the delays in samples (Δd_{12} and Δd_{13}) can be seen clearly. Shot 3 with the theoretical delays of -37 and -48 and the estimated delays of -32 and -49 demonstrates the minimum error of the series and highlights the system's potential. Underlying the positive improvement trend, the minimum and maximum errors follow a comparable declining pattern across all stages. The standard deviation indicates different behav-

ior between the two test series. On average, the standard deviation values are lower in the throwing test compared to the shooting test. This suggests that the frequency responses in the shooting test are more variable, likely due to greater differences in shot velocities. In contrast, the throwing test appears to produce more consistent velocities, resulting in lower variation and more stable localization errors. However, within the shooting series, the standard deviation shows a clear decreasing trend that aligns with the median values, indicating systematic improvement. In the throwing series, by contrast, the variation remains relatively constant across all stages.

Table 3.1.: Throw Series Statistics

| Metric | Stage 1 | Stage 2 | Stage 3 | Stage 4 |
|-------------------------|---------|---------|---------|---------|
| Size of Impact Array | 9 | 9 | 9 | 9 |
| Minimum Error / cm | 61 | 16 | 30 | 5 |
| Maximum Error / cm | 116 | 114 | 91 | 64 |
| Median Error / cm | 105 | 62 | 52 | 37 |
| Standard Deviation / cm | 16.41 | 34.35 | 20.05 | 20.04 |

Table 3.2.: Shot Series Statistics

| Metric | Stage 1 | Stage 2 | Stage 3 | Stage 4 |
|-------------------------|---------|---------|---------|---------|
| Size of Impact Array | 7 | 12 | 12 | 12 |
| Minimum Error / cm | 20 | 6 | 15 | 3 |
| Maximum Error / cm | 135 | 128 | 99 | 77 |
| Median Error / cm | 75 | 71 | 57 | 23 |
| Standard Deviation / cm | 44.48 | 39.79 | 22.87 | 22.59 |

3.2. Spatial Error Distribution

The scatter plot (see Figure 3.2) shows a direct comparison between actual and estimated positions for the worst (Stage 1) and best (Stage 4) test result. Each vector represents the offset between real and an estimated impact position. In Stage 1 (dashed blue lines), vectors are longer and inconsistently directed, reflecting a lack of precision in spatial resolution. In contrast, Stage 4 (full green lines) shows shorter vectors, demonstrating that the final system iteration can localize impacts closer to their true positions. This reinforces the effectiveness of the improvements implemented throughout the development process.

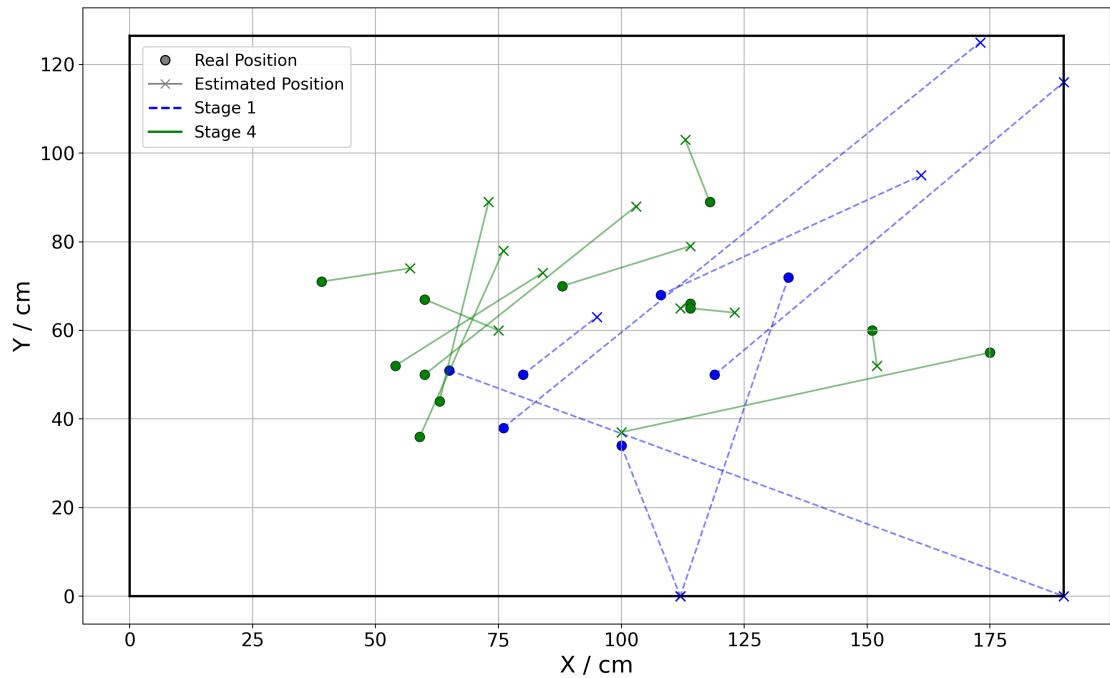


Figure 3.2.: The vector lengths visualizing error offset in Stage 1 and Stage 4.

4. Discussion

4.1. Evaluation of System Development Stages

The results indicate a clear trend of improving accuracy over the system's development (3.2), underlining the increasing reliability of the method. However, despite these advancements in stage 4, the median localization error of 23 cm remains too high to justify the current stage of the system for effective and practical SSL in the goal wall application.

In stage 1, the undamped acoustic setup combined with a rigid windowing strategy applied only to the reference microphone 1 results in excessive signal reflections and poor alignment across channels. The TDOA estimation is unreliable, and localization accuracy suffers enormously due to echoes and noise and the nonfunctional cross correlation. As seen in figure 4.1, differentiating between echoes and the actual impact is challenging. The system is likely to detect incorrect peaks, which leads to false delay estimations and strong drifts in the TDOA calculation. As a result, the localization error increases significantly. The tables A.2 and A.4 offer the comparison of the theoretical impact coordinates and their delays with the estimated delays. The estimated delays of shot 7 are -339 and -386 samples, where the correct ones lie at 55 and -44 samples. These metrics quantify the magnitude of the resulting error and underlining the space of optimization.

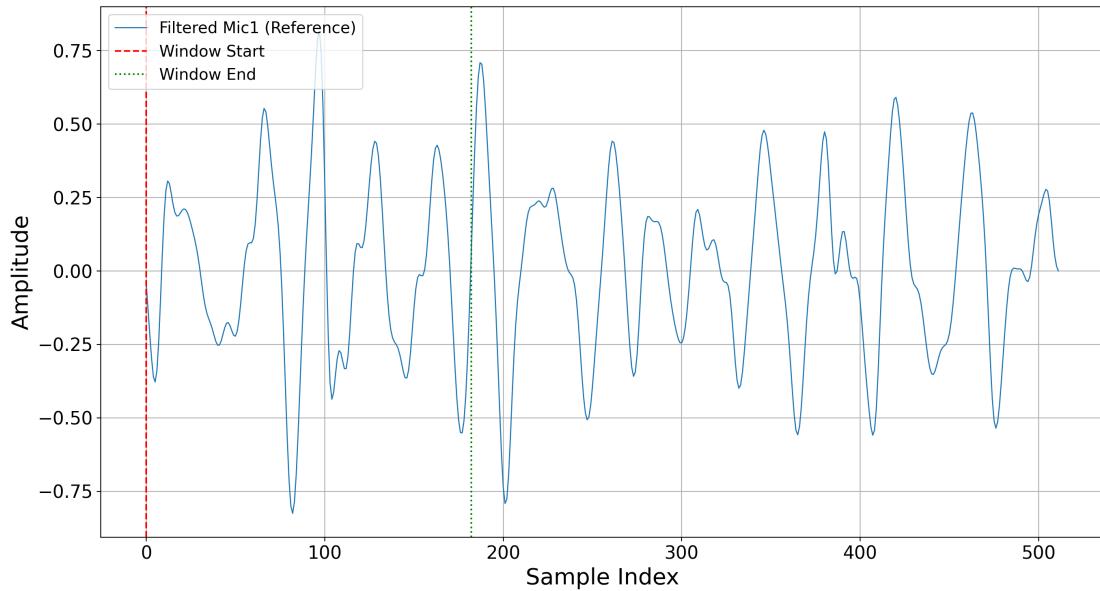


Figure 4.1.: Windowing in stage 1

Stage 2 introduces physical damping to minimize echoing and reflections of the surrounding. While this reduces the energy of late reflections, the improvement is just slightly because the windowing strategy remains unchanged and cross-correlation is applied to the full signal lengths of microphones 2 and 3. As a result, strong echo peaks are falsely identified as time shifts, leading to incorrect delay estimations. The individual windowing, is then applied in stage 3, where all microphone signals are individually windowed using a fixed-length frame (186 samples), which defines the maximum possible delay of the wall size with the microphone configuration. This strategy excludes known secondary echo peaks (e.g around sample 333) and thereby helps to isolate the initial impulse response more effectively. As a result, delay estimation improves and median errors begin to decrease, although the method still lacks to signal variability across different impact scenarios. Finally, Stage 4 combines a revised microphone configuration with optimized spatial placement and an enhanced windowing strategy. Figure 4.2 visualizes the cross correlation after the window is now set to the first relevant peaks and already excludes high echoing peaks. As a result high echo peaks, such as at sample -250, are not detected as a delay. This change improves spatial resolution, impact peak detection and results in better delay estimation. However, ongoing issues like the microphones' AGC, variations in individual signal quality, and remaining reflections continued to bring up errors. Although theoretical delay inputs yield accurate positional reconstruction, real-world signal conditions still bring up significant challenges, indicating that further algorithmic robustness is required before the system can be considered as practically reliable.

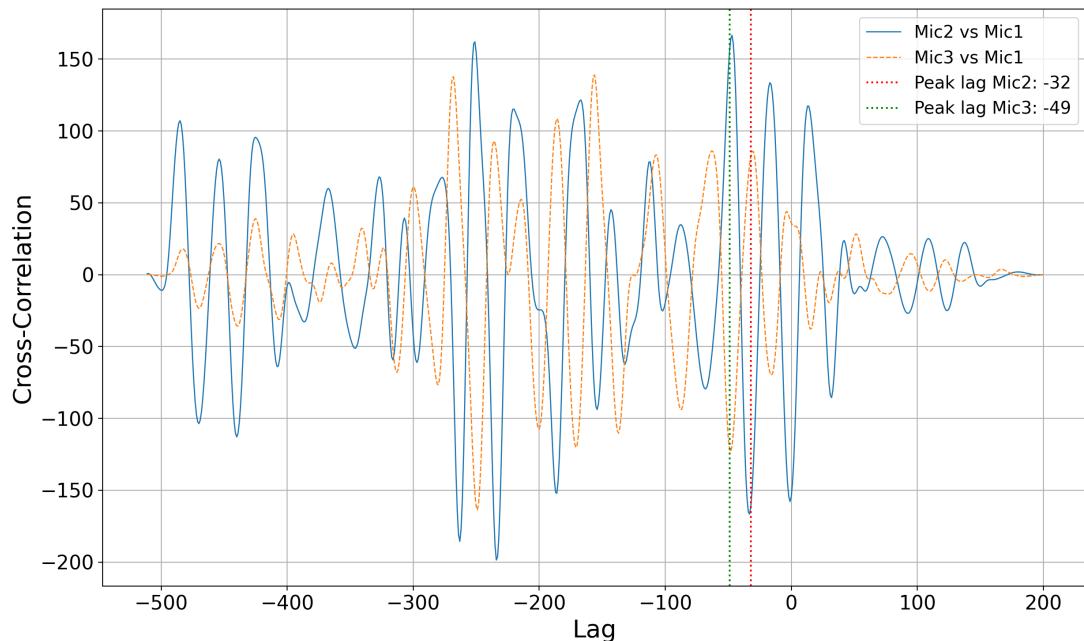


Figure 4.2.: Cross correlation in stage 4

Figure 4.2 gives an example of functional windowing. However, the main challenge remains detecting the exact moment of impact to determine the correct delays constantly in every variant of shot or throw, regardless to their velocity and resulting frequency. The position calculation itself works well, as proven when using the theoretical delays running the current software. However, since the system often struggles to detect the correct delays accurately from the acoustic signals, improvements to both the hardware and signal processing are still needed.

4.2. Future Improvements and Comparison to State of the Art

In comparison to already existing real world applications and SLL approaches the used methods and materials in this project might not be the most suitable. Firstly, a clean recording of the raw signal can significantly reduce the risk of errors during signal processing and have impact to the optimization of the outcome. Instead of the used analog MAX9814 microphone board, MEMS microphones without AGC limits prevent detected plateaus and their reparation. Unlike the microphones used in this project, real-world gunshot detection systems typically rely on MEMS-based directional microphones. Their mechanical architecture allows omnidirectional and high-frequency sensitivity without relying on AGC or high post-processing. The use of capacitive comb structures improves precision in detecting sound wave direction based on diaphragm displacement and resonant frequency [31].

One of the main problems in the current system is finding the correct delay between the microphone signals. Methods like cross-correlation can work, but they often fail when the signals have multiple peaks due to noise or echoes like it is given in the tested prototype. Deep learning-based approaches may estimate the time delay between two audio signals more accurately. A neural network takes the normalized cross correlation between two signals and learns to predict the delay directly. This technique shows better results than traditional methods, especially in environments with strong reflections and low signal quality [32]. Using a similar approach in this system could improve the accuracy of impact localization, especially in complex acoustic settings like an ice hockey arena.

Another field for improvement is the software configuration, particularly the sampling frequency. A higher sampling rate increases the time resolution (see Table 4.1), allowing for more accurate estimation of time delays between microphones. This is especially important in TDOA-based acoustic localization systems, where precise delay measurements are critical for accurate position estimation [29]. Currently the sampling frequency is set to 44 100 Hz but it can be increased up to 96 000 Hz using the Teensy 4.1 microcontroller.

Table 4.1.: Comparison of different of sampling frequency magnitudes regarding to sample and spatial resolution.

| f_s | resolution per sample | spatial resolution |
|----------|-----------------------|--------------------|
| 44100 Hz | 22.68 μ s | 7.78 mm |
| 96000 Hz | 10.42 μ s | 3.57 mm |

In addition, the ADC settings play a crucial role in determining signal quality during sampling. Although higher conversion speeds, such as the *VERY_HIGH_SPEED* mode, offer faster sampling rates, they also tend to introduce more electronic noise into the measurement. This trade-off is critical in SSL, where clear signals directly affect the time delay estimation. Increasing the ADC clock results in higher noise levels, which potentially compromises the resolution of the signal [33]. For the future, adjusting the ADC speed to achieve a balance between precision and responsiveness is essential for improving system accuracy.

Given the challenges of a real-world environment such as an ice hockey arena, facing lots of background noise, strong echoes and strong reflective surfaces, the system's future usability depends on its ability to operate reliably under these complex conditions. Even though these environments are acoustically complex, the tested approach still offers great potential. Traditional TDOA techniques based on cross-correlation may suffer from inaccurate peak detection due to reflections and overlapping noise sources, but to overcome these limitations, a promising alternative could be the use of machine learning models such as convolutional neural networks (CNNs) [18]. These models are capable of learning and differing between the individual acoustic patterns across the variety of frequencies due to different velocities and reflecting effects. For the ice hockey

goal wall application, a CNN could be trained to analyze the synchronized microphone signals and directly predict the location of puck impacts, eliminating the need for manually defined windowing or sensitive thresholding also eliminates their resulting errors. This approach would allow the wooden goal wall to serve as a reliable impact detection surface even under more complex environmental settings. With a sufficient dataset of labeled impacts, such a system could be trained to perform efficient and more accurate even in a real-time localization for ice hockey puck impacts on the wooden wall.

5. Conclusion

This thesis demonstrates the feasibility of an acoustic localization system using the hybrid of TDOA and AOA algorithms. The goal is to test the approach of locating the sound impact of an ice hockey puck hitting a wooden wall mounted along the goal line. Therefore a required hardware framework is developed to attach a wooden wall at a standardized goal frame. A triangular microphone array of three receivers captures data of two testing series in four different stages of the developed prototype. Over these stages, the acoustic approach can be evaluated for its reliability of use for the application. Every single development stage is quantified and compared with their localization error during the tests. While the error at stage 1 has a magnitude of 75 cm at the shooting series, it decreases over all the development iterations to 23 cm at stage 4. The system aims to be further developed and offer ice hockey players a digital training device where they can improve their timing and shot accuracy supported by smart technology. While the current system shows a promising trend, several possible aspects for future work remain. Improving signal quality through hardware adaptions, optimized processing settings, and enhance the time delay estimation technique could bring the current accuracy to a real world working application. From a broader perspective, the system's low cost and minimal hardware complexity make it a economic alternative to camera based solutions, especially in resource limited and low cost environments. Future iterations could further expand the system and include machine learning for a may more robust peak detection and even event classification.

With continued development, this acoustic localization system could evolve into a flexible, modular platform tailored specifically for ice hockey training. By integrating it with interactive elements such as LED targets or adaptive mini-games, it holds potential for use in precision shooting, reaction training, and performance feedback. Therefore the localization algorithms offer the fundamental base. With continued improvements in the future, a smart ice hockey goal wall could become part of the growing field of sports technology. Together with established systems like Hawk-Eye, Video Assistant referee, or Smart Wearables, supporting objective analysis and enhancing training environments.

Bibliography

- [1] Y. Wang, “The impact of digital transformation in the sports industry,” *Advances in Economics, Management and Political Sciences*, vol. 77, pp. 1–6, 4 2024.
- [2] S. A. Evans, “The biomechanics of ice hockey: Health and performance using wearable technology,” *Journal of Men’s Health*, vol. 18, p. 1, 9 2022.
- [3] N. Sieber, S. Walser, T. Weber, R. Gubler, H. Badertscher, and P. Eggenberger, “Development and evaluation of an ai-based exergame training system for ice-hockey players: a randomized controlled trial,” *Current Issues in Sport Science (CISS)*, vol. 10, p. 017, 1 2025.
- [4] T. Léger, P. J. Renaud, S. M. Robbins, and D. J. Pearsall, “Pilot study of embedded imu sensors and machine learning algorithms for automated ice hockey stick fitting,” *Sensors*, vol. 22, p. 3419, 4 2022.
- [5] T. Søderholm, “‘smart’ sports equipment and wearables have the potential to hugely impact the future of sports tracking and performance analytics,” Nordic Semiconductor, Tech. Rep., 2024. [Online]. Available: <https://www.nordicsemi.com/Nordic-news/2020/08/The-future-of-sports-tracking-is-in-the-equipment>
- [6] M. Labayen, I. G. Olaizola, N. Aginako, and J. Florez, “Accurate ball trajectory tracking and 3d visualization for computer-assisted sports broadcast,” Tech. Rep. [Online]. Available: <http://virtualeye.tv/>
- [7] L. M. Jayalath, “Hawk eye technology used in cricket,” *South Asian Research Journal of Engineering and Technology*, vol. 3, pp. 55–67, 4 2021.
- [8] A. Fitzpatrick, J. A. Stone, S. Choppin, and J. Kelley, “Analysing hawk-eye ball-tracking data to explore successful serving and returning strategies at wimbledon,” *International Journal of Performance Analysis in Sport*, vol. 24, pp. 251–268, 5 2024.
- [9] A. Ishfaque and B. Kim, “Real-time sound source localization in robots using fly *ormia ochracea* inspired mems directional microphone,” *IEEE Sensors Letters*, vol. 7, pp. 1–4, 1 2023.
- [10] X. Li, J. Chen, W. Qi, and R. Zhou, “A distributed sound source surveillance system using autonomous vehicle network,” in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 5 2018, pp. 42–46.

- [11] J. Chen, R. Takashima, X. Guo, Z. Zhang, X. Xu, T. Takiguchi, and E. R. Hancock, “Multimodal fusion for indoor sound source localization,” *Pattern Recognition*, vol. 115, p. 107906, 7 2021.
- [12] L. Petrica, “An evaluation of low-power microphone array sound source localization for deforestation detection,” *Applied Acoustics*, vol. 113, pp. 162–169, 12 2016.
- [13] M. Yilmaz and B. Gunel, “Investigation of acoustic source localization algorithm performances for gunshot localization on helicopters,” in *2016 24th Signal Processing and Communication Application Conference (SIU)*. IEEE, 5 2016, pp. 1885–1888.
- [14] N. S. Syahira, I. S. A. Halim, S. L. M. Hassan, and W. F. H. Abdullah, “Design and performance analysis of sound source localization using time difference of arrival estimation,” *Journal of Advanced Research in Applied Mechanics*, vol. 106, pp. 14–26, 6 2023.
- [15] M. A. Chung, H. C. Chou, and C. W. Lin, “Sound localization based on acoustic source using multiple microphone array in an indoor environment,” *Electronics (Switzerland)*, vol. 11, 3 2022.
- [16] M. U. Liaquat, H. S. Munawar, A. Rahman, Z. Qadir, A. Z. Kouzani, and M. A. Mahmud, “Localization of sound sources: A systematic review,” 7 2021.
- [17] X. Kang, D. Wang, Y. Shao, M. Ma, and T. Zhang, “An efficient hybrid multi-station tdoa and single-station aoa localization method,” *IEEE Transactions on Wireless Communications*, vol. 22, pp. 5657–5670, 8 2023.
- [18] L. Cao, H. Chen, Y. Chen, Y. Yue, and X. Zhang, “Bio-inspired swarm intelligence optimization algorithm-aided hybrid tdoa/aoa-based localization,” *Biomimetics*, vol. 8, p. 186, 4 2023.
- [19] J. A. Travieso-Rodriguez, R. Jerez-Mesa, J. Llumà, O. Traver-Ramos, G. Gomez-Gras, and J. J. R. Rovira, “Mechanical properties of 3d-printing polylactic acid parts subjected to bending stress and fatigue testing,” *Materials*, vol. 12, p. 3859, 11 2019.
- [20] M. A. Shah, I. A. Shah, D. G. Lee, and S. Hur, “Design approaches of mems microphones for enhanced performance,” 2019.
- [21] S. A. Zawawi, A. A. Hamzah, B. Y. Majlis, and F. Mohd-Yasin, “A review of mems capacitive microphones,” 5 2020.
- [22] “Max9814 - microphone amplifier with agc and low-noise microphone bias,” ANALOG DEVICES, Tech. Rep., 2022. [Online]. Available: www.analog.com
- [23] N. Kumar and A. Singh, “Study of microphone array characteristics and noise reduction,” Tech. Rep., 2018. [Online]. Available: <http://www.ripublication.com>
- [24] R. Michon, D. Overholt, S. Letz, Y. Orlarey, D. Fober, and C. Dumitrascu, “A Faust Architecture for the ESP32 Microcontroller,” in *Sound and Music*

- Computing Conference (SMC-20)*, Turin, Italy, Jun. 2020. [Online]. Available: <https://hal.science/hal-02988312>
- [25] J. Tomarakos, "Interfacing i2s-compatible audio devices to the adsp-21065l serial ports," 2018.
- [26] PJRC Forum User, "Adc library with support for teensy 4, 3.x and lc," <https://forum.pjrc.com/index.php?threads/adc-library-with-support-for-teensy-4-3-x-and-lc.25532/>, 2014, accessed: 2025-06-28.
- [27] laguna christopher and lerch alexander, "an efficient algorithm for clipping detection and declipping audio," *journal of the audio engineering society*, no. 9682, september 2016.
- [28] O. M. S. Jr, "Psd computations using welch's method. [power spectral density (psd)]," Sandia National Laboratories (SNL), Tech. Rep., 12 1991.
- [29] N. Zhu and T. Reza, "A modified cross-correlation algorithm to achieve the time difference of arrival in sound source localization," *Measurement and Control*, vol. 52, pp. 212–221, 3 2019.
- [30] X. Alameda-Pineda and R. Horaud, "A geometric approach to sound source localization from time-delay estimates," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 1082–1095, 6 2014.
- [31] J. Namitha and B. A. Thomas, "Design and simulation of mems based directional microphone for gun firing detection," *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) NCESC*, 2018.
- [32] Z.-Q. Wang, X. Zhang, and D. Wang, "Robust tdoa estimation based on time-frequency masking and deep neural networks." in *Interspeech*, 2018, pp. 322–326.
- [33] Texas Instruments, *Achieving Better Analog-to-Digital Converter Performance: Practical Techniques for Design Engineers*, 2015, application Report SLYT595. [Online]. Available: <https://www.ti.com/lit/an/slyt595/slyt595.pdf>

List of Figures

| | |
|--|----|
| 2.1. Clamp Design | 5 |
| 2.2. Frame and wall mounting setup | 6 |
| 2.3. Microphone mounting | 8 |
| 2.4. Microphone Configuration | 9 |
| 2.5. Teensy 4.1 | 10 |
| 2.6. Teensy mounted configuration | 10 |
| 2.7. Flowchart Signal Processing | 12 |
| 2.8. Raw microphone signals | 13 |
| 2.9. Repaired microphone signals | 14 |
| 2.10. Welch's method | 15 |
| 2.11. Bandpass filtered signal | 15 |
| 2.12. Windowed signal | 16 |
| 2.13. Cross correlation 1/2 | 17 |
| 2.14. Cross Correlation 1/3 | 17 |
| 2.15. Heatmap | 19 |
| 2.16. Target positions and Impact mark | 20 |
| 2.17. Shielded microphones at stage 2 | 21 |
| 2.18. Windowing Stage 4 | 23 |
| 3.1. Boxplots of Testing Series | 25 |
| 3.2. Scatter Plot Shot Series | 27 |
| 4.1. Windowing in stage 1 | 29 |
| 4.2. Cross correlation in stage 4 | 30 |

List of Tables

| | |
|---|------|
| 2.1. Pin Setting MAX9814 | 8 |
| 2.2. Impact of Time Delay | 12 |
| 2.3. Positions Throw Series | 19 |
| 2.4. Microphone Configuration Stage 4 | 22 |
| 3.1. Throw Series Statistics | 26 |
| 3.2. Shot Series Statistics | 26 |
| 4.1. Comparsion of sampling frequencies | 31 |
| A.1. Theoretical Throw Metrics of stage 1 | XII |
| A.2. Theoretical Shot Metrics of Stage 1 | XII |
| A.3. Throw Metrics of Stage 1 | XIII |
| A.4. Shot Metrics of Stage 1 | XIII |
| A.5. Theoretical Throw Metrics of Stage 4 | XIII |
| A.6. Theoretical Shot Metrics of Stage 4 | XIV |
| A.7. Throw Metrics of Stage 4 | XIV |
| A.8. Shot Metrics of Stage 4 | XIV |

List of Symbols

| symbol | name | unit |
|------------------|----------------------------|------------------------------|
| c | speed of sound | $\frac{\text{cm}}{\text{s}}$ |
| f_s | sampling frequency | Hz |
| x, y, z | spatial coordinates | cm |
| r_{21}, r_{31} | distance difference | cm |
| $tdoaij$ | time difference of arrival | s |
| lag | sample shift | samples |
| Δt | delay in samples | samples |
| θ | angle of arrival | rad |
| σ | angle uncertainty | rad |
| d_{ij} | distance | cm |
| $error$ | localization error | cm |

Abbreviations

| | |
|------|------------------------------|
| SSL | Sound Source Localization |
| AOA | Angle of Arrival |
| TDOA | Time Difference of Arrival |
| GPS | Global Positioning System |
| TOA | Time of Arrival |
| DOA | Direction of Arrival |
| PLA | Polyactide |
| CAD | Computer Aided Design |
| STL | Standard Triangle Language |
| AGC | Automatic Gain Control |
| I2S | Inter-IC Sound Bus |
| CSV | Comma Separated Values |
| ADC | Analog Digital Converter |
| FFT | Fast Fourier Transformation |
| CNN | Convolutional Neural Network |

A. GitHub repository and full metrics tables

The following link provides access to the GitHub repository containing all code and data used in this thesis: https://github.com/bastisee/Bachelorthesis_Seelos

Table A.1.: Theoretical Throw Metrics of stage 1

| Point | (x, y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg |
|-------|------------|-----------------|-----------------|-----------|-----------|
| P1 | (30, 110) | 163 | 131 | -0.8176 | -46.85 |
| P2 | (90, 110) | 13 | 32 | -0.1180 | -6.76 |
| P3 | (150, 110) | -139 | -26 | -0.0627 | -3.59 |
| P4 | (30, 80) | 134 | 66 | -1.3759 | -78.84 |
| P5 | (90, 80) | 11 | -19 | -0.5127 | -29.38 |
| P6 | (150, 80) | -117 | -64 | -0.2897 | -16.60 |
| P7 | (30, 40) | 102 | -12 | -1.4749 | -84.51 |
| P8 | (90, 40) | 9 | -103 | -0.8574 | -49.13 |
| P9 | (150, 40) | -89 | -116 | -0.5490 | -31.46 |

Table A.2.: Theoretical Shot Metrics of Stage 1

| Shot | (x/y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg |
|------|-----------|-----------------|-----------------|-----------|-----------|
| #2 | (80, 50) | 28 | -68 | -0.8689 | -49.78 |
| #5 | (76, 38) | 33 | -88 | -0.9814 | -56.23 |
| #7 | (65, 51) | 55 | -44 | -1.0055 | -57.61 |
| #8 | (100, 34) | -8 | -125 | -0.8256 | -47.30 |
| #9 | (119, 50) | -44 | -100 | -0.6138 | -35.17 |
| #10 | (108, 68) | -27 | -60 | -0.5292 | -30.32 |
| #11 | (134, 72) | -82 | -69 | -0.3913 | -22.42 |

Table A.3.: Throw Metrics of Stage 1

| Point | (x, y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg | Estimated (x/y) | Error / cm |
|-------|------------|-----------------|-----------------|-----------|-----------|-----------------|------------|
| P1 | (30, 110) | -48 | -89 | -1,5386 | -88,15 | 120/56 | 105 |
| P2 | (90, 110) | -42 | -152 | -2,2122 | -126,75 | 125/9 | 106 |
| P3 | (150, 110) | -49 | -155 | -2,2831 | -130,00 | 132/0 | 111 |
| P4 | (30, 80) | -47 | -333 | -2,5167 | -144,19 | 114/0 | 116 |
| P5 | (90, 80) | -189 | -32 | -1,4181 | -65,78 | 178/118 | 96 |
| P6 | (150, 80) | -129 | -174 | -2,5167 | -144,19 | 190/0 | 89 |
| P7 | (30, 40) | -54 | -33 | -1,5450 | -66,15 | 117/88 | 101 |
| P8 | (90, 40) | -159 | -273 | -2,5167 | -144,19 | 190/0 | 108 |
| P9 | (150, 40) | 39 | -229 | -2,5167 | -144,19 | 95/14 | 61 |

Table A.4.: Shot Metrics of Stage 1

| Shot | (x/y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg | Estimated (x/y) | Error / cm |
|------|--------|-----------------|-----------------|-----------|-----------|-----------------|------------|
| #2 | 80/50 | -1 | -69 | -1,3252 | -75,93 | 95/63 | 20 |
| #5 | 76/38 | -231 | 21 | -0,8137 | -46,62 | 173/125 | 131 |
| #7 | 65/51 | -339 | -386 | -2,5167 | -144,19 | 190/0 | 135 |
| #8 | 100/34 | -29 | -240 | -2,5167 | -144,19 | 112/0 | 36 |
| #9 | 119/50 | -272 | -45 | -1,2322 | -70,6 | 190/116 | 97 |
| #10 | 108/68 | -150 | -49 | -1,2585 | -72,1 | 161/95 | 60 |
| #11 | 134/72 | -32 | -268 | -2,5167 | -144,19 | 112/0 | 75 |

Table A.5.: Theoretical Throw Metrics of Stage 4

| Point | (x, y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg |
|-------|------------|-----------------|-----------------|-----------|-----------|
| P1 | (30, 110) | 151 | 138 | -1.5708 | 90.00 |
| P2 | (90, 110) | 13 | 52 | -0.1244 | -7.13 |
| P3 | (150, 110) | -133 | -6 | -0.0624 | -3.58 |
| P4 | (30, 80) | 125 | 75 | -1.5708 | -90.00 |
| P5 | (90, 80) | 12 | 1 | -0.5586 | -32.01 |
| P6 | (150, 80) | -110 | -45 | -0.3029 | -17.35 |
| P7 | (30, 40) | 95 | -7 | -1.5708 | -90.00 |
| P8 | (90, 40) | 9 | -83 | -0.9120 | -52.25 |
| P9 | (150, 40) | -83 | -101 | -0.5734 | -32.86 |

Table A.6.: Theoretical Shot Metrics of Stage 4

| Shot | (x/y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg |
|------|--------|-----------------|-----------------|-----------|-----------|
| #1 | 63/44 | 53 | -41 | -1,1488 | -65,82 |
| #3 | 114/66 | -37 | -48 | -0,55 | -31,51 |
| #4 | 59/36 | 56 | -52 | -1,2289 | -70,41 |
| #5 | 88/70 | 15 | -17 | -0,6862 | -39,32 |
| #6 | 151/60 | -97 | -74 | -0,4436 | -25,42 |
| #7 | 175/55 | -120 | -84 | -0,407 | -23,32 |
| #8 | 54/52 | 71 | -13 | -1,2196 | -69,88 |
| #9 | 60/50 | 60 | -25 | -1,1526 | -66,04 |
| #11 | 60/67 | 68 | 12 | -1,0348 | -59,29 |
| #14 | 118/89 | -52 | -11 | -0,3132 | -17,95 |
| #15 | 114/65 | -37 | -50 | -0,5586 | -32,01 |
| #16 | 39/71 | 105 | 46 | -1,3796 | -79,05 |

Table A.7.: Throw Metrics of Stage 4

| Point | (x, y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg | Estimated (x/y) | Error / cm |
|-------|------------|-----------------|-----------------|-----------|-----------|-----------------|------------|
| P1 | (30, 110) | 0 | 86 | -0,4779 | -27,38 | 92/126 | 64 |
| P2 | (90, 110) | 21 | 28 | -0,8606 | -49,31 | 86/92 | 18 |
| P3 | (150, 110) | -126 | -14 | -1,1119 | -64,11 | 149/103 | 7 |
| P4 | (30, 80) | 66 | 49 | -0,7281 | -41,72 | 66/89 | 37 |
| P5 | (90, 80) | -42 | -18 | -1,1436 | -65,52 | 114/83 | 24 |
| P6 | (150, 80) | -95 | 11 | -0,9656 | -55,33 | 133/114 | 38 |
| P7 | (30, 40) | 61 | -4 | n.s | n.s | 63/61 | 39 |
| P8 | (90, 40) | 16 | -79 | -1,5385 | -88,15 | 85/40 | 5 |
| P9 | (150, 40) | 10 | -59 | -1,403 | -80,39 | 90/51 | 61 |

Table A.8.: Shot Metrics of Stage 4

| Shot | (x/y) | Δd_{21} | Δd_{31} | AOA / rad | AOA / deg | Estimated (x/y) | Error / cm |
|------|--------|-----------------|-----------------|-----------|-----------|-----------------|------------|
| #1 | 63/44 | 52 | 40 | -0,7854 | -45 | 73/89 | 45 |
| #3 | 114/66 | -32 | -49 | -1,3381 | -76,67 | 112/65 | 3 |
| #4 | 59/36 | 41 | 14 | -0,9472 | -54,27 | 76/78 | 45 |
| #5 | 88/70 | -40 | -25 | -1,1869 | -68 | 114/79 | 27 |
| #6 | 151/60 | -93 | -85 | -1,581 | -90,59 | 152/52 | 8 |
| #7 | 175/55 | -8 | -97 | -1,6697 | -95,67 | 100/37 | 77 |
| #8 | 54/52 | 23 | -7 | n.s | n.s | 84/73 | 37 |
| #9 | 60/50 | -18 | 1 | n.s | n.s | 103/88 | 57 |
| #11 | 60/67 | 38 | -22 | -1,1683 | -66,94 | 75/60 | 17 |
| #14 | 118/89 | -43 | 16 | -0,9349 | -53,56 | 113/103 | 15 |
| #15 | 114/65 | -53 | -58 | -1,3965 | -80,01 | 123/64 | 9 |
| #16 | 39/71 | 77 | 31 | -0,8419 | -48,24 | 57/74 | 19 |